

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein -- protein search, using sw model

Run on: April 13, 2006, 12:51:03 ; Search time 189 Seconds  
(without alignments)  
683.478 Million cell updates/sec

Title: US-10-686-490d-2  
Perfect score: 1509  
Sequence: 1 MAQVKANGITLYEEQGRH.....ERHPLRIELIAGHRAAAEA 294

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_21.\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*  
9: Geneseqp2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1509	100.0	294	ADO26338	Ado26338 Anti-Kasl
2	607.5	40.3	322	ADA34452	Ada34452 Acinetoba
3	597	39.6	299	ADRS1257	Adrs1257 Anti-biof
4	589.5	39.1	321	ABB09186	Abb09186 Acinetoba
5	586.5	38.9	312	ABB09185	Abb09185 Acinetoba
6	562.5	37.3	422	ABO73253	AbO73253 Pseudomon
7	483	32.0	360	ADL05923	Adl05923 M. catarr
8	314.5	20.8	267	AAW30522	Aaw30522 Kurthia s
9	262.5	17.4	247	ABM95684	Abm95684 M. xanthu
10	227.5	15.1	271	ADI23937	Adi23937 Streptomy
11	220.5	14.6	276	ABU21130	Abu21130 Protein e
12	216.5	14.3	264	ADI23904	Adi23904 Streptomy
13	209.5	13.9	259	AAAG6464	Aag64664 Hydroxyl
14	209.5	13.9	259	ADO70988	Ado70988 Stenotro
15	204	13.5	271	AAO22154	Aao22154 Ramoplani
16	203.5	13.5	284	AEB39939	Aeb39939 L. pneumo
17	203.5	13.5	280	AEB36545	Aeb36545 L. pneumo
18	198	13.1	232	AAE24160	Aae24160 Human 466
19	198	13.1	232	ADD12925	Adt12925 Human alp
20	198	13.1	232	ADT05828	Adt05828 Hidden Ma
21	198	13.1	232	ADT05777	Adt05777 Hidden Ma
22	196	13.0	283	AAAI0975	Aari0975 Polynhydro
23	192	12.7	287	ABO73720	AbO73720 Pseudomon
24	189.5	12.6	251	AAAB60302	Aab60302 Alpha/bet

25	189	12.5	268	7	ABO83496	AbO83496 Pseudomon
26	185.5	12.3	233	7	ADB80253	Adb80253 Human 588
27	185	12.3	265	9	ABM92440	Abm92440 M. xanthu
28	184.5	12.2	325	3	AAG13760	Aag13760 Arabidops
29	184.5	12.2	647	5	ABB91674	Abb91674 Arabidops
30	184.5	12.2	647	8	ADN73561	Adn73561 Thale cre
31	183.5	12.2	251	4	AA879868	Aab79868 Corynebac
32	183.5	12.2	251	4	AAG92372	Aag92372 C glucam
33	183.5	12.2	251	7	ADD13483	Add13483 C. glutam
34	183	12.1	289	2	AAW99184	Aaw99184 Rhodococc
35	182.5	12.1	271	9	ABM93099	Abm93099 M. xanthu
36	182	12.1	614	5	ABB91334	Abb91334 Arabidops
37	181.5	12.0	264	5	ABP40747	Abp40747 Staphyloc
38	181.5	12.0	264	8	ADS08054	AdS08054 Staphyloc
39	180.5	12.0	258	6	ABU42585	Abu42585 Protein e
40	180	11.9	338	7	ABO83828	AbO83828 Pseudomon
41	175.5	11.6	289	8	ADR13741	Adr13741 Amidase,
42	175.5	11.6	339	6	ABU21279	Abu21279 Protein e
43	175	11.6	261	9	ABM94555	Abm94555 M. xanthu
44	174.5	11.6	396	3	AAG48813	Aag48813 Arabidops
45	174.5	11.6	396	3	AAG48791	Aag48791 Arabidops

ALIGNMENTS

RESULT 1  
ADO26338  
ID ADO26338 standard; protein; 294 AA.  
XX  
AC ADO26338;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Anti-Kaslauskas lipase protein.  
XX  
KW ds; gene; enzyme; anti-Kaslauskas lipase; esterification; hydrolysis;  
KW pharmaceutical; agricultural chemical; liquid crystal compound.  
XX  
OS Synthetic.  
XX  
PN EP1418237-A2.  
XX  
PD 12-MAY-2004.  
XX  
PF 06-OCT-2003; 2003EP-00022590.  
XX  
PR 16-OCT-2002; 2002DE-01048166.  
XX  
(FARB ) BAYER CHEM AG.  
XX  
Bosch B, Meissner R, Berendes F, Koch R;  
WPI; 2004-378759/36.  
N-PSDB; ADO26337.  
XX  
New nucleic acid encoding anti-Kaslauskas lipase and derived enzymes,  
PT useful for stereospecific hydrolysis and synthesis of aralkyl esters,  
intermediates for pharmaceuticals and plant protection agents.  
XX  
Claim 11; Page 16-17; 22pp; German.  
XX  
The present invention provides the protein and coding sequences of an  
anti-Kaslauskas lipase. These are used as catalysts in esterification or  
hydrolysis reactions for preparation of enantiomeric aralkanol, or their  
esters, useful for preparing pharmaceuticals or agricultural chemicals,  
also liquid crystal compounds. The present sequence is the protein of the  
invention.  
XX  
SQ Sequence 294 AA;  
Query Match 100.0%; Score 1509; DB 8; Length 294;  
Best Local Similarity 100.0%; Pred. No. 1e-151;

Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQVKANGITLEYEQHRRHPSMLIMGLGGQLIDWPPEFIRGLAERGFRVICFDRDA 60  
 DB 1 MAQVKANGITLEYEQHRRHPSMLIMGLGGQLIDWPPEFIRGLAERGFRVICFDRDA 60  
 QY 61 GLSTKLEGVKPNIARVFLASMLKPRVPYTTDDMALDTVGLMDALGIESHTVVGVSMG 120  
 DB 61 GLSTKLEGVKPNIARVFLASMLKPRVPYTTDDMALDTVGLMDALGIESHTVVGVSMG 120  
 QY 121 GMIAQLGAKHGERVKSITLMTSSGNPRMPAPQVLOKFMVPSKMDKEWIKYNLEL 180  
 DB 121 GMIAQLGAKHGERVKSITLMTSSGNPRMPAPQVLOKFMVPSKMDKEWIKYNLEL 180  
 QY 181 LTTIGSPGLDREKALDVRKSIERCICPEGTORQLAAILOSQSRVKLLRIAVPTLVISG 240  
 DB 181 LTTIGSPGLDREKALDVRKSIERCICPEGTORQLAAILOSQSRVKLLRIAVPTLVISG 240  
 QY 241 AEDPLPYOCGRDIADHIPGARFELIEGMDHPIPERHIPRLIELIAGHAAAAEA 294  
 DB 241 AEDPLPYOCGRDIADHIPGARFELIEGMDHPIPERHIPRLIELIAGHAAAAEA 294

RESULT 2  
 ADA34452  
 ID ADA34452 standard; protein; 322 AA.  
 AC ADA34452;  
 DT 20-NOV-2003 (first entry)  
 DE Acinetobacter baumannii protein #1613.  
 KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;  
 KW plant biocontrol agent.  
 OS Acinetobacter baumannii.  
 PN US6562958-B1.  
 XX 13-MAY-2003.  
 XX 04-JUN-1999; 99US-00328352.  
 PR 09-JUN-1998; 98US-0088701P.  
 PA (GENO-) GENOME THERAPEUTICS CORP.  
 PI Breton G, Bush D;  
 DR WPI; 2003-576092/54.  
 DR N-PSDB; ADA30326.

New Acinetobacter baumannii proteins and nucleic acids, useful as reagents for diagnosing a bacterial disease, as components of antibacterial vaccines, as targets for antibacterial drugs, or as biocontrol agents for plants.  
 Example; SEQ ID NO 5739; 328pp; English.  
 The invention relates to isolated Acinetobacter baumannii nucleic acids. The A. baumannii nucleic acids and polypeptides are useful as reagents for diagnosing a bacterial disease, as components of antibacterial vaccines, as targets for antibacterial drugs, to detect the presence of A. baumannii and other Acinetobacter species in a sample, in screening compounds for the ability to interfere with the A. baumannii life cycle or to inhibit A. baumannii infection, and as biocontrol agents for plants. The present sequence represents the amino acid sequence of an A. baumannii protein.  
 Sequence 322 AA;  
 Query Match 40.3%; Score 607.5; DB 6; Length 322;

Best Local Similarity 46.9%; Pred. No. 1.4e-55;  
 Matches 134; Conservative 49; Mismatches 98; Indels 5; Gaps 4;

QY 7 NGITLEYEQHRRHPSMLIMGLGGQLIDWPPEFIRGLAERGFRVICFDRDAGLSTKL 66  
 DB 33 NGIELHVEVGKPEPTILLIMGLGAQMLFPDFFCKSLIDQGFVRVIRPDNRDILGLSSKV 92  
 QY 67 --EGVKKPNIARVFLASMLKPR--VPYTTDDMALDTVGLMDALGIESHTVVGVSMGMI 123  
 DB 93 RHQG-KRLNTWKMGKRFALGURNQCAPYTLVDMADDVSMLLDRLGVSXKAVLGASMGMI 151  
 QY 124 AQILGAKHGERVKSITLMTSSGNPRMPAPQVLOKFMVPSKMDKEWIKYNLELTT 183  
 DB 152 AQILAAKYPEKVEKGLMFTNNQFPFLPPPKQLLSLIGKPESDRDESGIVNHSKLFL 211  
 QY 184 IGSPG-LDREKALDVRKSIERCICPEGTORQLAAILOSQSRVKLLRIAVPTLVISGAB 242  
 DB 212 IGSPGYINHIEAVOTARKLYQRSYYPAGVLOQFLAILCTGSLLOLDREIKQPTLVHGSR 271  
 QY 243 DPLLPYOCGRDIADHIPGARFELIEGMDHPIPERHIPRLIELIAGH 288  
 DB 272 DRLLPPSHGKAVAKAISGAKFELIDGMGHDIPAHFIPQLSGLFAHH 317

RESULT 3  
 ADR51257  
 ID ADR51257 standard; protein; 299 AA.  
 XX AC ADR51257;  
 DT 04-NOV-2004 (first entry)  
 DE Anti-biofilm polypeptide #24.  
 KW antibacterial; antiinflammatory; anti-biofilm polypeptide;  
 KW humoral response; environmental sample; water; liquid; soil; air;  
 KW biological sample; medical device; pharmaceutical; food product;  
 KW cosmetic; hygiene product; water treatment device; water transport; pulp;  
 KW paper processing; paper recycling equipment; toothpaste; chewing gum;  
 KW mouthwash; dental cleaner; caries; plaque; gingivitis; periodontitis;  
 KW oropharyngeal colonization; gastric colonization; thermostability;  
 KW thermotolerance.  
 XX OS Unidentified.  
 XX FH Key Location/Qualifiers  
 FT Domain 49.288  
 FT /note= "alpha/beta hydrolase fold domain"  
 XX PN WO2004066945-A2.  
 XX PD 12-AUG-2004.  
 XX PF 26-JAN-2004; 2004WO-US002242.  
 XX PR 24-JAN-2003; 2003US-0442794P.  
 XX PA (DIVE-) DIVERSA CORP.  
 XX PI Barton N, Robertson D, Chang K, Elkins J;  
 DR WPI; 2004-625302/60.  
 DR N-PSDB; ADR51256.  
 XX New anti-biofilm enzymes such as esterase, glycosidase, amylase, useful for controlling biofilms on a wide range of household, industrial, medical surfaces.  
 XX Claim 45; SEQ ID NO 48; 413pp; English.  
 CC The invention relates to an isolated or recombinant anti-biofilm polypeptide (I) having at least 50-100% sequence identity to 58 fully defined sequences (S1), over region of at least 100 residues, where the

sequence identities is determined by analysis with a sequence comparison algorithm or by visual inspection. (I) is useful for making an antibody by administering to a non-human animal (I) or a DNA (II) encoding (I), to generate a humoral response. Oligonucleotide primers (III) are useful for isolating or recovering (II) encoding (I) from an environmental sample comprising water, liquid, soil, air or biological sample derived from a bacterial, protozoan, insect, yeast, plant, fungal or mammalian cell. (I) is useful for making a small molecule by providing several biosynthetic (I) capable of synthesizing or modifying a small molecule, and reacting the substrate with enzymes. (I) is useful for preventing the growth of a biofilm on a medical device, pharmaceutical, food product, device for making a food, cosmetic, hygiene product, water treatment device, water transport or storage device or pulp and paper processing and paper recycling equipment. The pharmaceutical is a tablet, pill, implant, suppository, inhaler, spray or ointment. The biofilm is from Pseudomonas or Staphylococcus, which involves administering to the biofilm (I) having esterase activity. (I) is useful in cleaning and decontaminating hard surfaces such as floors, equipment, machinery and industrial water treatments. (I) is useful as preservatives in food, medicinal, hygiene, and cosmetic products, useful in toothpaste, chewing gum, mouthwashes, dental cleaners. (I) is useful for controlling dental biofilms associated with caries, plaque, gingivitis, and periodontitis. (I) is useful for treating oropharyngeal and gastric colonization by pathogenic microorganisms. The activity of (I) is thermostable or thermotolerant. This sequence corresponds to a polypeptide of the invention.

Sequence 299 AA;

Query Match 39.6%; Score 597; DB 8; Length 299;  
Best Local Similarity 45.1%; Pred. No. 1.7e-54;  
Matches 128; Conservative 43; Mismatches 111; Indels 2; Gaps 1;  
QY 9 ITLEYEQHRRHPSMLLMGLGQQLIDWPEEFIRGLAERGFVCFDNRDAGLSTKLEG 68  
DB 8 IREYETFGHPDDPAIVLMGLGQQLILWPEAFCRMADAGHVYVRFNDRDGLSTHLDH 67  
QY 69 VKPNRIARVFLASMGKRPVPTLDDMALDVTGLMDALGISTHVVGVSGMGMAIILG 128  
DB 68 LRPNPLAALROALPLRVASVYTLDDMADVAGLLDNLNIQAHVGVSGMGMAIQLA 127  
QY 129 AKHGERVKSITLMTSSGNPRMPAPRQVLOKFMVPEKMDKEWIKYNLELLTTIGSPG 188  
DB 128 ARHATKVRSLTLLMTSSGNPRMPAPRQVLOKFMVPEKMDKEWIKYNLELLTTIGSPG 187  
QY 189 LDREKIALD--VRKSIERCLCEPTQORQALAIQSGSRVKKLRRIAVPTLVISGAEPL 246  
DB 188 YPKPEALRIVAEGDFRAPHFAGFNRQLHVAALAPSRAPLPRIKQPADVIHGADLIV 247  
QY 247 PYOCGRDIADHIPGAFELIEGMGHDIPERHPIRLIELIAGHAA 290  
DB 248 PVAARDLVRRLPNATLIDVPGMGHDFPTEIMPRIARRIVETAA 291

RESULT 4  
ABB09186  
ID ABB09186 standard; protein; 321 AA.

AC ABB09186;  
XX 04-JUL-2002 (first entry)

DE Acinetobacter calcoaceticus carboxyl esterase #2.

KW Pseudomonas aeruginosa; esterase; estA; enzyme; enantiomeric;  
KW optical active carboxylic acid; asymmetric hydrolysis; hypotensive;  
KW carboxylic ester racemate; hypertension; analapril; captopril.

OS Acinetobacter calcoaceticus.

XX WO200132847-A1.

XX 10-MAY-2001.

XX

PF 01-NOV-2000; 2000WO-KR001243.  
XX 01-NOV-1999; 99KR-00047927.  
PR (KORE-) KOREA RES INST BIOSCIENCE & BIOTECHNOLOG.  
XX (BIOH-) BIOHOLDINGS INC.  
XX Kim C, Rhee S, Song K, Lee J, Boyapaty G;  
PI WPI; 2001-316432/33.  
XX Novel esterase derived from Pseudomonas aeruginosa useful for producing  
PT optical active carboxylic acids and pharmaceuticals, preferably  
PT hypertension treatments drugs like analapril or captopril.  
XX Example 6; Fig 3; 40pp; English.  
XX The present invention describes an esterase (I) derived from Pseudomonas  
CC aeruginosa which has a sequence comprising 315 amino acids, and produces  
CC optical active carboxylic acids and enantiomeric isomers by the  
CC asymmetric hydrolysis of the carboxylic ester racemates. (I) has  
CC hypotensive activity. (I) is useful for producing optical active  
CC carboxylic acids, preferably optical active (R)-carboxylic acids from  
CC carboxylic acid ester racemates. (I) is useful for the production of  
CC various kinds of physiologically active pharmaceuticals, specially  
CC hypertension treatments drugs like analapril or captopril. (I) has  
CC excellent ability for producing optically active carboxylic acids from  
CC carboxylic acid esters. Production of optically active carboxylic acids  
CC by (I) is highly more selective, simple and environmentally favorable  
CC when compared to conventional techniques. (I) is thermally stable at high  
CC temperatures above 70 degrees Celsius. The present sequence represents a  
CC carboxyl esterase amino acid sequence given in comparison with the  
CC esterase of the present invention  
XX  
SQ Sequence 321 AA;

Query Match 39.1%; Score 589.5; DB 4; Length 321;  
Best Local Similarity 45.5%; Pred. No. 1.2e-53;  
Matches 130; Conservative 53; Mismatches 100; Indels 3; Gaps 3;

QY 6 ANGTLEYEQHRRHPSMLLMGLGQQLIDWPEEFIRGLAERGFVCFDNRDAGLSTK 65  
DB 33 SNGLEHVEVGGNPDHPHTILLMGLGAQMLFWDFFCKSLIDQGFVIRFDRDGLSSK 92  
QY 66 LEGV-KKPNRIARVFLASMGK-LKRPVPTLDDMALDVTGLMDALGISTHVVGVSGMGMI 123  
DB 93 IRRHKGRLNTLKLMSRFTLGLNGOAGPYTLVDMADVSLLEAMRIKKVNVICASMGMI 152  
QY 124 AQILGAKHGERVKSITLMTSSGNPRMPAPRQVLOKFMVPEKMDKEWIKYNLELLTT 183  
DB 153 AQITAAKYPEKVEKALMFTSNQPLLPPLPPPPKQFLSIGKPKSDEDOGIINHSUKLFEI 212  
QY 184 IGSPGLDREKIALD--VRKSIERCLCEPTQORQALAIQSGSRVKKLRRIAVPTLVISGAE 242  
DB 213 IGSPGVNQVEAIQTARKLYORSYHPAGVLOQFLAILCTGSLQLDQKISQTLVHGRS 272  
QY 243 DPLLPYOCGRDIADHIPGAFELIEGMGHDIPERHPIRLIELIAGH 288  
DB 273 DRLLPPSHGKAVAKAIGKAFELIQMGHDIPPHFIPQLSYLFAHH 318

RESULT 5  
ABB09185  
ID ABB09185 standard; protein; 312 AA.

XX ABB09185;

XX 04-JUL-2002 (first entry)

DE Acinetobacter calcoaceticus carboxyl esterase #1.

XX Pseudomonas aeruginosa; esterase; estA; enzyme; enantiomeric;  
KW optical active carboxylic acid; asymmetric hydrolysis; hypotensive;

KW carboxylic ester racemate; hypertension; analapril; captopril.

XX Acinetobacter calcoaceticus.

XX WO200132847-A1.

XX 10-MAY-2001.

XX 01-NOV-2000; 2000WO-KR001243.

XX 01-NOV-1999; 99KR-00047927.

XX (KORE-) KOREA RES INST BIOSCIENCE & BIOTECHNOLOG.

XX (BIOH-) BIOHOLDINGS INC.

XX Kim C, Rhee S, Song K, Lee J, Boyapaty G;

XX WPI; 2001-316432/33.

XX Novel esterase derived from Pseudomonas aeruginosa useful for producing

XX optical active carboxylic acids and pharmaceuticals, preferably

XX hypertension treatments drugs like analapril or captopril.

XX Example 6; Fig 3; 40pp; English.

XX The present invention describes an esterase (I) derived from Pseudomonas  
XX aeruginosa which has a sequence comprising 315 amino acids, and produces  
XX optical active carboxylic acids and enantiomeric isomers by the  
XX asymmetric hydrolysis of the carboxylic ester racemates. (I) has  
XX hypotensive activity. (I) is useful for producing optical active  
XX carboxylic acids, preferably optical active (R)-carboxylic acids from  
XX carboxylic acid ester racemates. (I) is useful for the production of  
XX various kinds of physiologically active pharmaceuticals, specially  
XX hypertension treatments drugs like analapril or captopril. (I) has  
XX excellent ability for producing optically active carboxylic acids from  
XX carboxylic acid esters. Production of optically active carboxylic acids  
XX by (I) is highly more selective, simple and environmentally favourable  
XX when compared to conventional techniques. (I) is thermally stable at high  
XX temperatures above 70 degrees Celsius. The present sequence represents a  
XX carboxyl esterase amino acid sequence given in comparison with the  
XX esterase of the present invention

XX Sequence 312 AA;

Query Match 38.9%; Score 586.5; DB 4; Length 312;

Best Local Similarity 45.1%; Pred. No. 2.3e-53;

Matches 129; Conservative 54; Mismatches 100; Indels 3; Gaps 3;

QY 6 ANGITLEYEQGHRHHPHMLLMGLGQLDWPPEEFIRGLAERGFVFCFDRDAGLSTK 65

DB 24 SNGLELHVEVGGPDHPTILLIIMGLGQMLFWDFFCKSLIDGQFYVIRFDNRDGLSSK 83

QY 66 LEGV-KKPNIAARVFLASMG-L-KRPVPTLDDMALDTVGLMDALGIESTHVGVSGMMI 123

DB 84 IRHKGRKRLNTKLMSRFTLGLNGQAPYTYLDMAEDVSLLEAMRKVNIVIGASGMMI 143

QY 124 AQLGAKHGERVKSITLMTITSSGNRPMAPRPQVLQKFMVRPKSMDEKWKYNLELLTT 183

DB 144 AQTIAATPEKVKALMTFSSNNQPLPPFPKQPLSLIGKPKSDEDDGLIINSLKLF 203

QY 184 IGSPGLDREKALD-VKRSERCLCEGTORQALAIQSGSRVKLLRRIAVPTLVISGAE 242

DB 204 IGSPGVNVEATQARLLYQSRVHPAGVLQOQFLAICTGSLQLDQKLSQPTLVHGRS 263

QY 243 DPLPYOCGRDIAHDPGARFELIEGMDHIDPERHIPRLIELIAGH 288

DB 264 DRLPPSHGKAVAKAIGKARFELIQGMGHDIPPHFIPQLSYLPFAH 309

RESULT 6

ABO73253

ID ABO73253 standard; protein; 422 AA.

XX

AC ABO73253;

XX 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #5428.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

OS US6551795-B1.

PN 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

DR N-PSDB; ABD06824.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
XX useful as molecular targets for diagnostics, prophylaxis and treatment of  
XX pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 21999; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
XX polynucleotides encoding them. The sequences are useful in diagnosis and  
XX therapy of pathological conditions, as molecular targets for diagnostics,  
XX prophylaxis and treatment of pathological conditions resulting from a  
XX bacterial infection, for evaluating a compound, such as a polypeptide,  
XX for the ability to bind a P. aeruginosa nucleic acid, as components of  
XX effective antibacterial targets, as targets for antibacterial drugs,  
XX including anti-P. aeruginosa drugs, as templates for recombinant  
XX production of P. aeruginosa-derived peptides or polypeptides, as target  
XX components for diagnosis and/or treatment of P. aeruginosa-caused  
XX infection, and in detection of P. aeruginosa sequences or other sequences  
XX of Pseudomonas species using biochip technology. Sequences ABO67826-  
XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
XX sequence data for this patent did not form part of the printed  
XX specification but was obtained in electronic format from USPTO at  
XX seqdata.uspto.gov/sequence.html

XX Sequence 422 AA;

Query Match 37.3%; Score 562.5; DB 7; Length 422;

Best Local Similarity 43.1%; Pred. No. 1.3e-50;

Matches 125; Conservative 45; Mismatches 117; Indels 3; Gaps 2;

QY 4 VKANGITLEYEQGHRHHPHMLLMGLGQLDWPPEEFIRGLAERGFVFCFDRDAGLS 63

DB 125 VELGDVRLAYQSIGRSDPALLIVMGLGQLIHDPVVVSALCEQGFVIRYDNRDVGLS 184

QY 64 TKLEGVKKPNIAARVFLASMG-L-KRPVPTLDDMALDTVGLMDALGIESTHVGVSGMMI 123

DB 185 ANNVFVPSRLLTYEVVVRVRLGLPVSAPYTLTDMAGDALHLLDALDIPQARVLGASGMMI 244

QY 124 AQLGAKHGERVKSITLMTITSSGNRPMAPRPQVLQKFMVRPKSMDEKWKYNLELLTT 183

DB 245 AQHIADMAPQRLTSLTTLVTSSGASGLPAPSSLLRLRLAR-REASREAOVEQADLLAA 303

QY 184 IGSPGL--DREKALDVRKSIERCLCEGTORQALAIQSGSRVKLLRRIAVPTLVISGA 241

DB 304 LGSPEVRDDRQQLLQAAARSYDRAFNPEGVQRLAILAEPKRVPLNRLQVPTLVHGT 363

QY 242 EDPLPYOCGRDIAHDPGARFELIEGMDHIDPERHIPRLIELIAGHAA 291

Db 364 ADPLLPVHGVHVAHNRGSELKLIPLGLAHRFQAEKPEKPIIAAVVPYLKA 413

RESULT 7

ADL05923  
ID ADL05923 standard; protein; 360 AA.

XX AC  
XX ADL05923;  
XX DT (first entry)

XX DE 06-MAY-2004 (first entry)

XX DE M. catarrhalis protein #1689.

XX KW Moraxella catarrhalis; infection.

XX OS Moraxella catarrhalis.

XX PN US6673910-B1.

XX PD 06-JAN-2004.

XX PF 04-APR-2000; 2000US-00540236.

XX PR 08-APR-1999; 99US-0128416P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX PI Breton GL;

XX DR WPI; 2004-178127/17.

XX DR N-PSDB; ADL04003.

XX PT New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for preparing a composition for diagnosing, preventing or treating infection caused by Moraxella catarrhalis.

XX PS Disclosure; SEQ ID NO 3609; 429pp; English.

XX CC The invention relates to an isolated nucleic acid encoding an Moraxella catarrhalis polypeptide. The nucleic acid is useful for preparing a composition for diagnosing, preventing or treating infection caused by Moraxella catarrhalis. The present sequence represents the amino acid sequence of a M. catarrhalis protein.

XX SQ Sequence 360 AA;

Query Match 32.08; Score 483; DB 8; Length 360;  
Best Local Similarity 37.34; Pred. No. 3.1e-42;  
Matches 107; Conservative 61; Mismatches 113; Indels 6; Gaps 3;

QY 8 GITLEVEEEOGHRHPSMLLMGLGGQLIDWPPEFIRGLAERGFVFCFNDRLAGLSTK-- 65

Db 42 GIKLVEAGNPEHPMIFITGLSQMFHSDQFLKPFIDAGFVIFPNDRTGLSKIQ 101

QY 66 LEGYKFNPIARVFLASMGKLPK---VPYTLDDMALDTVGLMDALGISTHSTHVGVSGMG 122

Db 102 IDGLPRLNTFKMLQKQAGLSNRSEPVAYTLTDMEDARLITKMLHNVNLIGASWGM 161

QY 123 IAQILGNHGERVKSGLTMTTSSGNPNPAPROVLOKFNVPKSMDSKKEWIKYNLELT 182

Db 162 IAQIVARYPKYIKNLVLLFSTNSRAFLRPNPKQFMTFVRPESHESRDMVRHSWNFMT 221

QY 183 TIGSPG-LDREKALDVKRSIERCLCEGTQORLAAILQSGSRVKLLRIAVPTLVISGA 241

Db 222 AVGSPGHLDTKGTFAIAEKYQNFHPLXVQQUTAILASRSILRTKQIRANTLVHGN 281

QY 242 EDPLPYCCGRDIADHIPGARFELIEGMDHIDIPERHPIELIELIAGH 288

Db 282 KDGIVAPNHGKILAKVISNARFVLVDGMGHDLPNYYYPYINGLISEH 328

RESULT 8

AAW30522

AAW30522 standard; protein; 267 AA.

XX AC  
XX AAW30522;

XX DT 17-OCT-2003 (revised)

XX DT 26-OCT-1998 (first entry)

XX DE Kurthia sp. bioH gene-encoded polypeptide.

XX KW Biotin; vitamin; bioH gene; pimelyl CoA.

XX OS Kurthia sp. 538-KA26; (DSM 10609).

XX PN EP853127-A2.

XX PD 15-JUL-1998.

XX PF 18-SEP-1997; 97EP-00116237.

XX PR 27-SEP-1996; 96EP-00115540.

XX PA (HOFF ) HOFFMANN LA ROCHE & CO AG F.

XX PI Furuichi Y, Hoshino T, Kimura H, Kiyasu T, Nagahashi Y;

XX DR WPI; 1998-364652/32.

XX DR N-PSDB; AAV42058.

XX PT New DNA and vectors encoding polypeptides - used for recombinant production of biotin.

XX PS Claim 1; Page 24-25; 45pp; English.

XX CC This is a polypeptide encoded by a newly isolated bioH gene (see AAV42058) of Kurthia sp. 538-KA26 (DSM 10609). This polypeptide is thought to be involved in the synthesis of the biotin precursor pimelyl CoA. Polypeptides (see AAW30518-25) encoded by novel bioD, bioA, bioP, bioB, bioH, bioFII, bioHII and bioC genes (see AAV42054-61) of Kurthia sp. are provided, as well as vectors comprising one or more of the genes, transformed cells, and a process for the production of biotin that comprises cultivating the transformed cells and isolating biotin from the culture medium. Biotin is used in the preparation of pharmaceutical, food or feed compositions. The synthesis is an improvement on prior art methods involving fermentation of e.g. E.coli and Bacillus sphaericus, which have low productivity due to accumulation of DTB, a biotin precursor. (Updated on 17-OCT-2003 to standardise OS field)

XX SQ Sequence 267 AA;

Query Match 20.84; Score 314.5; DB 2; Length 267;  
Best Local Similarity 33.14; Pred. No. 1.8e-24;  
Matches 92; Conservative 43; Mismatches 104; Indels 39; Gaps 9;

QY 1 MAQYKANGITLVEEEOGHRHPSMLLMGLGGQLIDWPPEFIRGLAERGFVFCFNDRLA 60

Db 1 MPFVNHNESLYEV--HGQGDPLLLINGLVNLSW-HRTVFTLAKR-FKIVFDNRGV 56

QY 61 GLSTKLEGYKFNPIARVFLASMGKLPKPRVPTLDDMALDTVGLMDALGISTHSTHVGVSGM 120

Db 57 GKSS-----KPEQPSYIEMMAEDARAVLDVAVSDSAHYVIGSMG 95

QY 121 GMIAQLGAKHGERVKSGLTMTTSSGNPNPAPROV---LQKFMVPKSMDSKKEW---- 173

Db 96 GMIAQRLAITYPERSVSLVGLGCTTAGTTHIQSPSISTLMVSRASLTGSPRDNALAAP 155

QY 174 IKYNLELLTIGSPGLDREKALDVKRSIERCLCEGTQORLAAILQSGSRVKLLRIAV 233

Db 156 IVYSQAFIEK--HPELIQ-----DIQRIEITPPSAYLSQLOACLTHDTSNE-LDKINI 208

QY 234 PTLVTSIGAEDPLLPYOCGRDIADHIPGARFELIEGMDH 271

Db 209 PTLIHGDADNLVPYENGKMLAERIOGSGFHTVSCAGH 246

```

RESULT 9
ABM95684
ID ABM95684 standard; protein; 247 AA.
XX AC ABM95684;
XX DT 02-JUN-2005 (first entry)
XX DE M. xanthus protein sequence, seq id 14883.
XX KW Transgenic plant; DNA replication; gene regulation; gene expression.
XX OS Myxococcus xanthus.
XX PN US6833447-B1.
XX PD 21-DEC-2004.
XX PF 10-JUL-2001; 2001US-00902540.
XX PR 10-JUL-2000; 2000US-0217883P.
XX PA (MONS ) MONSANTO TECHNOLOGY LLC.
XX PI Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;
XX WPI; 2005-028716/03.
XX PT New substantially purified Myxococcus xanthus nucleic acid molecule
PT encoding a nitrite reductase, useful for determining gene expression,
PT identifying mutations in a gene of interest, and for constructing
PT mutations in a gene of interest.
XX PS Example 2; SEQ ID NO 14883; 25pp; English.
XX CC The invention relates to a substantially purified nucleic acid molecule
CC encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
CC recombinant DNA construct for expression of a nitrite reductase gene in a
CC plant cell, and a plant cell comprising the recombinant DNA construct.
CC The nucleic acid is useful for determining gene expression, identifying
CC mutations in a gene of interest, and for constructing mutations in a gene
CC of interest. Sequences given in records for SEQ IDs 9692-16825 represent
CC a group of 7134 Myxococcus xanthus proteins and peptides. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
XX SQ Sequence 247 AA;
Query Match 17.4%; Score 262.5; DB 9; Length 247;
Best Local Similarity 30.3%; Pred. No. 5.7e-19;
Matches 79; Conservative 38; Mismatches 107; Indels 37; Gaps 7;
QY 45 LAERGFVFCFNRDAGLSTKLEGVKKPNIARVFLLASMGKPRVPYTLDDMALDVTGLM 104
DB 4 LARAGRFVIRVDHRTGSR-----SHGVPGATPTTLDLAAADVISVL 45
QY 105 DALGIETHVGVSMGMIQILGAKHGERVKSITLM---ITSGNPRMPAPRVOVKPF 161
DB 46 DGYGIERAHLVMSGLGLLQCVALKYPERVLSITLISAQIFSEPPDFGMDPAVLAHF 105
QY 162 MRVP--KSMDEKWKYNLEI-LTTIGSP--GLDREKALDVRKSIERCLCPEGTOROLA 216
DB 106 ORAATLNSDEAEAGFQVLESRLCVGRARRSFDEARVARAVQDYRRALAPQ-----C 159
QY 217 AILQSGSRVKL-----LRRIAVPTLVISGAEDPLLPYQCGRDIADHIPGARFELIEGMGH 271
DB 160 ALNHAGLSGLGMYGRTREIEAPLLVINGSDVPVIDHAHGVALSRVAKGACLVTLHDAGH 219
QY 272 DIPERHPIRLIELIAGHAAA 292
DB 220 DLHPDDWETWTRAITAHTSAA 240

```

---

```

RESULT 10
ADI23937
ID ADI23937 standard; protein; 271 AA.
XX AC ADI23937;
XX DT 22-APR-2004 (first entry)
XX DE Streptomyces refuineus 024A locus ORF9 protein.
XX KW antimicrobial; fungicide; virucide; gene therapy; lipopeptide synthesis;
XX AS4145; NRRL 3143; antimicrobial; antifungal; antiviral;
XX biosynthetic locus; 024A; ORF9.
XX OS Streptomyces refuineus.
XX PN US2003198981-A1.
XX PD 23-OCT-2003.
XX PF 24-DEC-2002; 2002US-00329079.
XX PR 26-DEC-2001; 2001US-0342133P.
XX PR 17-APR-2002; 2002US-0372789P.
XX PR 03-SEP-2002; 2002US-00232370.
XX PA (ECOP-) ECOPIA BIOSCIENCES INC.
XX PI Farnet CM, Staffa A, Zazopoulos E;
XX WPI; 2003-852784/79.
XX DR N-PSDB; ADI23938.
XX PT New isolated, purified or enriched nucleic acid, useful for synthesizing
PT lipopeptides, particularly from the biosynthetic locus AS4145 and NRRL
PT 3143, that exhibits antimicrobial, antifungal or antiviral activity.
XX PS Example 4; SEQ ID NO 51; 69pp; English.
XX CC The invention describes an isolated, purified or enriched nucleic acid
CC (i) comprising: any of 4 fully defined sequences (SEQ. ID NO: 1, 6, 17
CC and 34), and their coding regions; a nucleic acid having at least 75%
CC sequence identity to a nucleic acid of (a); or complements of (a) or (b).
CC The methods and compositions of the present invention are useful for
CC synthesizing lipopeptides, particularly AS4145 and NRRL 3143, exhibiting
CC antimicrobial, antifungal or antiviral activity. This is the amino acid
CC sequence of a protein encoded by the Streptomyces refuineus 024A (or NRRL
CC 3143) biosynthetic locus ORF9.
XX SQ Sequence 271 AA;
Query Match 15.1%; Score 227.5; DB 7; Length 271;
Best Local Similarity 26.9%; Pred. No. 3.5e-15;
Matches 84; Conservative 47; Mismatches 98; Indels 83; Gaps 13;
QY 1 MAQVKANGITLVEEGHHRHPSMLLIGLGGQLIDWPEEFIRGLAERGFVFCFNRDA 60
DB 1.MPTTRINGALDHRGTG--SGPVLIMSGSAKSAWHLHOVPALVAEGEATFTNR-- 56
QY 61 GLSTKLEGVKKPNIARVFLLASMGKPR--VPYTLDDMALDVTGLMDALGIESTHVGV 117
DB 57 -----GVPPSGGPGFTLGDMAADTVGLIHLGIGPCAVVGM 93
QY 118 SMGMIATQILGAKHGERVKSITLMTSSGNPRMPA-----PRPVLQK 160
DB 94 SLGARVAREVARTPDLVSRCLVAPRARSDRRAACTAAETALADSGVTLPPRYRAVR 153
QY 161 FNR--VPKSM--DKB--EWKYNLELTTIGSPGLDREKALDVRKSIERCLCPEGTORQ 214
DB 154 AMQNLSPTLADDRQIADWLDV--LELAADG--PGL-----RTQLE 191
QY 215 LAAILQSGSRVKLLRRIAVPTLVISGAEDPLLPYQCGRDIADHIPGARFELIEGMGH--- 271

```









CC benzodioxole to sesamol. The Stenotrophomonas maltophilia bacteria  
CC contains a gene cluster (9320 nucleotide sequence given in the  
CC specification) that is involved in the conversion of 1,3-benzodioxole to  
CC sesamol. The method of the invention is useful for preparing sesamol  
CC which can be used as a synthetic raw material for pharmaceuticals such as  
CC antihypertensives. The present amino acid sequence represents a protein  
CC encoded by the Stenotrophomonas maltophilia sesamol production-related  
CC gene cluster of the invention.  
XX  
SQ Sequence 259 AA;  
  
Query Match 13.9%; Score 209.5; DB 8; Length 259;  
Best Local Similarity 28.9%; Pred. No. 2.7e-13;  
Matches 82; Conservative 36; Mismatches 109; Indels 57; Gaps 10;  
  
QY 4 VKANGITLEYEQGHRHPSMLLMGLGQLIDW-----PEEFIRGLAERFRVCEP 56  
DB 13 VVANGIRTNTHDAG--DGAFLVMIHSGFGVSANWRLTMPE-----LATR-FRVIAPD 64  
  
QY 57 NRDAGLSTKLEGVKKNPNARVFLASMGKLPVPYTLDDMALDTVGLMDALGISTHVVVG 116  
DB 65 MVGFGYSQREPIH-----YSLDTWVQVALLDLDIEQASVVG 104  
  
QY 117 VSMGGMIAQLGAKHGERVKSILMITSSGNPRMPAPRPOVLQKFRVPSMDKEWIKY 176  
DB 105 NSFSGAIALALAIHPKRVKRLVLM-----GSGVSPFITEGLDAVWGYPQSVB----- 153  
  
QY 177 NLELLATTIGSPGLDREKLDVRSKIERCLCEGTORQLAAIL-----QSG-----SRVKL 227  
DB 154 NMRALLDIFA--YDRKLNDLQAMRYKASIQGFQESFSMPFPAPRQNGVEMWASPIEQ 211  
  
QY 228 LRRIAVPTLVISGAEDPLPYQCGRIADHPIGAPFELIEGMGH 271  
DB 212 IRGIEHQTLLVHGREDKVIPLQNSYELLOAIQVLAQLHVFQKCGH 255

RESULT 15  
AAO22154  
ID AAO22154 standard; protein; 271 AA.  
XX  
AC AAO22154;  
XX  
DT 03-OCT-2002 (first entry)  
DE Ramoplanin biosynthetic ORF 9 protein.  
XX  
KW Ramoplanin; ramoplanin biosynthetic pathway polypeptide; antibiotic;  
KW biosynthesis gene cluster; bioengineering; peptide synthetase module;  
KW adenylation domain; hydroxyphenylglycine; HPG; antibiotic precursor;  
KW chlorinate; lipopeptide.  
XX  
OS Actinoplanes sp.  
XX  
PN WO200231155-A2.  
XX  
PD 18-APR-2002.  
XX  
XX 15-OCT-2001; 2001WO-CA001462.  
XX  
PR 13-OCT-2000; 2000US-0239924P.  
PR 12-APR-2001; 2001US-0283296P.  
PR 24-JUL-2001; 2001US-00910813.  
XX  
XX (ECOP-) ECOPIA BIOSCIENCES INC.  
PA Farnet CW, Zazopoulos E, Staffa A;  
PI  
XX  
XX WPI: 2002-435445/46.  
DR N-PSDB; AAL40781.  
XX  
XX Novel isolated ramoplanin biosynthetic pathway polypeptide useful for  
PT chemically modifying biological molecule that is a substrate for a  
PT polypeptide encoded by a ramoplanin biosynthesis gene cluster.

XX  
PS Claim 14; Page 146-147; 212pp; English.  
XX  
CC The invention relates to an isolated ramoplanin biosynthetic pathway  
CC polypeptide selected from a polypeptide of open reading frames (ORF) 1-  
CC 32. The isolated polypeptides are useful for chemically modifying a  
CC biological molecule that is a substrate for a polypeptide encoded by a  
CC ramoplanin biosynthesis gene cluster, by contacting the biological  
CC molecule with the isolated polypeptide, where the polypeptide chemically  
CC modifies the biological molecule. The method comprises contacting the  
CC biological molecule with at least two different polypeptides encoded by  
CC ramoplanin ORFs 1-31. The polypeptides are useful for directing the  
CC biosynthesis of the antibiotic ramoplanin in microorganisms. An isolated  
CC gene cluster comprising the ORFs is useful as a substrate for  
CC bioengineering of antibiotic structures. An isolated polypeptide or its  
CC encoding nucleic acid sequence is useful for generating derivatives of  
CC ramoplanin, for improving production or for producing variants of other  
CC antibiotics of the peptide class. The isolated polypeptides are useful  
CC for synthesis of ramoplanin in vivo or in vitro, as an adenylation domain  
CC in conjunction with other peptide synthetase modules and allowing the  
CC incorporation of Thr into a peptide antibiotic precursor, for modifying  
CC fatty acid structure and/or enhancing fatty acid incorporation into the  
CC peptide antibiotic structure, for production of an hydroxyphenylglycine  
CC (HPG)-containing peptide antibiotic, for enhancing secretion of  
CC ramoplanin or its variants and derivatives, for enhancing uptake of  
CC precursors for ramoplanin biosynthesis, for enhancing production of  
CC ramoplanin products or its variants or derivatives, to chlorinate HPG of  
CC a peptide antibiotic precursor, and for designing specific nucleotide  
CC probes and primers for identifying and isolating putative lipopeptide  
CC -producing microorganisms. This sequence represents one of the ORF  
CC proteins of the ramoplanin producing Actinoplanes sp. microorganism of  
CC the invention  
XX  
SQ Sequence 271 AA;  
  
Query Match 13.5%; Score 204; DB 5; Length 271;  
Best Local Similarity 25.0%; Pred. No. 1.1e-12;  
Matches 78; Conservative 38; Mismatches 120; Indels 76; Gaps 9;  
  
QY 4 VKANGITLEYEQGHRHPSMLLMGLGQLIDWPEEFIRGLAERFRVCFDNRDAGLS 63  
DB 9 VVTNGVRLAYRSGAGE--PVLIMINGSGSAGQTTVHTQTPALHEAGYSTVWFDSE-- 61  
  
QY 64 TKLEGVKKNPNARVFLASMGKLPVPYTLDDMALDTVGLMDALGISTHVVGVSGGMI 123  
DB 62 ----GIPPSDV-----PAGKYSLABMTADTRGLIEALDLDLAPCRIVGTSLGAMI 105  
  
QY 124 AQILGAKHGERVKSILMITSSGNPRMPAPRQ-----VLQKFRVPSMDK----- 170  
DB 106 AQELAVDPHELVRCAVLATLARPDAARAAQADIDLESVTLTPAAYEATAVFMFS 165  
  
QY 171 -----EEWIKYNLELLTTIGSPGLDREKLDVRSKIERCLCEGTORQLAAILQ 220  
DB 166 PATLNDVAVREWLDI-FELSGTGVSA-----QAWAEL- 199  
  
QY 221 SGRVKLLRRIAVPTLVISGAEDPLPYQCGRIADHPIGAPFELIEGMGH-----DIPER 276  
DB 200 TGDRAALRSVTAPCKRVISFADDLITPPLHAAEVAEIPDCDLVEISRCGHLGRLRPDA 259  
  
QY 277 HIPRLIELIAGH 288  
DB 260 VNAAILFELDSE 271

Search completed: April 13, 2006, 12:55:20  
Job time : 192 secs

Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Lladman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kag, A.; Larbiq, K.; Liu

```

.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A/Reference number: A82950; MUID:20437337; PMID:10984043
A/Accession: D83197
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-328 <STO>
A/Cross-references: UNIPROT:Q9HY40; UNIPARC:UPI000000CSA79; GB:AE004779; GB:AE004091; NID
A/Experimental source: strain PA01
C/Genetics:
A/Gene: PA3586
C/Superfamily: tropinesterase

Query Match 37.3%; Score 562.5; DB 2; Length 328;
Best Local Similarity 43.1%; Pred. No. 2.7e-39;
Matches 125; Conservative 45; Mismatches 117; Indels 3; Gaps 2;

QY 4 VKANGITLEYEQGRRHPSMLIMGLGQGLIDWPPEFIRGLAERGFVFCFNDAGLS 63
DB 31 VELGDVRLAQSGRSDPALLVWGLGGLIHPDVEVSALCEQGRVIRYDNRDGLS 90
QY 64 TKLEGVKKPNIRVFLASMLKPRVPTLDDMALDTVGLMDALGIESTHVGVSGMGI 123
DB 91 AMNVVPSSRLTYEVVRYLGLPVSAFYTLTDMAGDALHLLDALDIPQAHVGLASMGMI 150
QY 124 AOLKAGHGRVKSLLTMITSSGNPRMPAPRQVLOKFMVPSKMDKEWIKYNLELTT 183
DB 151 AOHADMAPORLLSLTVMTSSGAEGLPAPSESLLRLAR-REAAAREQAVEQADLLAA 209
QY 184 IGSPGL--DREKALDVRKSIERCLCEGTQORQLAAILQSGSRVKLLRRIAVPTLVISGA 241
DB 210 LGSPEVDRDQQLLQOARSYDRAFNEGVQORQLAILAESRVPVLLNRLQVTLVHGT 269
QY 242 EDPLLPYQCGRDIAHPIGARFELIEGMDHPIPERHIPRIELIAGHAAA 291
DB 270 ADPLLPVHGVHAAHIRGSELKIPGLAHRFOEAFKPELLAAVVPYLKA 319

RESULT 3
C70614
probable lipG protein - Mycobacterium tuberculosis (strain H37Rv)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: C70614
R;Cole, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: C70614
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-301 <COL>
A/Cross-references: UNIPROT:P96935; UNIPARC:UPI00000165242; GB:Z92772; GB:AL123456; NID:9
A/Experimental source: strain H37Rv
C/Genetics:
A/Gene: lipG
C/Superfamily: carboxyl esterase

Query Match 32.1%; Score 485; DB 2; Length 301;
Best Local Similarity 40.1%; Pred. No. 6.9e-33;
Matches 112; Conservative 50; Mismatches 111; Indels 6; Gaps 4;

QY 9 ITLEYEQGRRHPSMLIMGLGQGLIDWPPEFIRGLAERGFVFCFNDAGLSKLEG 68
DB 13 VKLYYEDMGDLDPVLLIMGLGAQMLLWTDFCARLVAKGLRIVRYDNRDGLSKYTER 72
QY 69 VKKPN--IARVFLASMLKPRVPTLDDMALDTVGLMDALGIESTHVGVSGMGI 126
DB 73 -HFGQPLATVRSWGLSPQAAYTLEDMAAALLDHLDVKAHVGVSGMGI 131

.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A/Reference number: A82950; MUID:20437337; PMID:10984043
A/Accession: D83197
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-328 <STO>
A/Cross-references: UNIPROT:Q9HY40; UNIPARC:UPI000000CSA79; GB:AE004779; GB:AE004091; NID
A/Experimental source: strain PA01
C/Genetics:
A/Gene: PA3586
C/Superfamily: tropinesterase

Query Match 37.3%; Score 562.5; DB 2; Length 328;
Best Local Similarity 43.1%; Pred. No. 2.7e-39;
Matches 125; Conservative 45; Mismatches 117; Indels 3; Gaps 2;

QY 4 VKANGITLEYEQGRRHPSMLIMGLGQGLIDWPPEFIRGLAERGFVFCFNDAGLS 63
DB 31 VELGDVRLAQSGRSDPALLVWGLGGLIHPDVEVSALCEQGRVIRYDNRDGLS 90
QY 64 TKLEGVKKPNIRVFLASMLKPRVPTLDDMALDTVGLMDALGIESTHVGVSGMGI 123
DB 91 AMNVVPSSRLTYEVVRYLGLPVSAFYTLTDMAGDALHLLDALDIPQAHVGLASMGMI 150
QY 124 AOLKAGHGRVKSLLTMITSSGNPRMPAPRQVLOKFMVPSKMDKEWIKYNLELTT 183
DB 151 AOHADMAPORLLSLTVMTSSGAEGLPAPSESLLRLAR-REAAAREQAVEQADLLAA 209
QY 184 IGSPGL--DREKALDVRKSIERCLCEGTQORQLAAILQSGSRVKLLRRIAVPTLVISGA 241
DB 210 LGSPEVDRDQQLLQOARSYDRAFNEGVQORQLAILAESRVPVLLNRLQVTLVHGT 269
QY 242 EDPLLPYQCGRDIAHPIGARFELIEGMDHPIPERHIPRIELIAGHAAA 291
DB 270 ADPLLPVHGVHAAHIRGSELKIPGLAHRFOEAFKPELLAAVVPYLKA 319

RESULT 3
C70614
probable lipG protein - Mycobacterium tuberculosis (strain H37Rv)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: C70614
R;Cole, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: C70614
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-301 <COL>
A/Cross-references: UNIPROT:P96935; UNIPARC:UPI00000165242; GB:Z92772; GB:AL123456; NID:9
A/Experimental source: strain H37Rv
C/Genetics:
A/Gene: lipG
C/Superfamily: carboxyl esterase

Query Match 32.1%; Score 485; DB 2; Length 301;
Best Local Similarity 40.1%; Pred. No. 6.9e-33;
Matches 112; Conservative 50; Mismatches 111; Indels 6; Gaps 4;

QY 9 ITLEYEQGRRHPSMLIMGLGQGLIDWPPEFIRGLAERGFVFCFNDAGLSKLEG 68
DB 13 VKLYYEDMGDLDPVLLIMGLGAQMLLWTDFCARLVAKGLRIVRYDNRDGLSKYTER 72
QY 69 VKKPN--IARVFLASMLKPRVPTLDDMALDTVGLMDALGIESTHVGVSGMGI 126
DB 73 -HFGQPLATVRSWGLSPQAAYTLEDMAAALLDHLDVKAHVGVSGMGI 131

```

```

QY 127 LGAKHGRVKSLLTMITSSGNPRMPAPRQVLOKFMVPSKMDKEWIKYN-LELLTTIG 185
DB 132 FAARFAQRKTLLAVIFSSNNHRLPAPPAPRALALLTGPPDPSPROVDVNAVRSKIIG 191
QY 186 SPG--LDREKALDVRKSIERCLCEGTQORQLAAILQSGSRVKLLRRIAVPTLVISGAED 243
DB 192 SPAYPIPEQVRAEAAESYDRNPHFWGIAQOQFSAILGSGSLRYDRRIVAPTIVHGRAD 251
QY 244 PLLPYOCGRDIAHPIGARFELIEGMDHPIPERHIPRLI 282
DB 252 KLMRPFGRGAVARAINGARLVLDGMDHDLPRQLMDRVI 290

RESULT 4
E87146
probable hydrolase [imported] - Mycobacterium leprae
C/Species: Mycobacterium leprae
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C/Accession: E87146
R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.;
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A/Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.;
A/Title: Massive gene decay in the leprosy bacillus.
A/Reference number: A86909; MUID:21128732; PMID:11234002
A/Accession: E87146
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-304 <STO>
A/Cross-references: UNIPROT:Q9CBK5; UNIPARC:UPI000000CE03; GB:AL450380; NID:gl3093568;
C/Genetics:
A/Gene: lipG
C/Superfamily: carboxyl esterase

Query Match 31.5%; Score 475.5; DB 2; Length 304;
Best Local Similarity 36.9%; Pred. No. 4.3e-32;
Matches 109; Conservative 54; Mismatches 117; Indels 15; Gaps 4;

QY 9 ITLEYEQGRRHPSMLIMGLGQGLIDWPPEFIRGLAERGFVFCFNDAGLSKLEG 68
DB 13 LKLYYEDMGNDVDPVLLIMGLGAQLVLRATFCEKLVAGQLRVRYDNRDGLSRTDS 72
QY 69 VKKP-----NIARVFLASMLKPRVPTLDDMALDTVGLMDALGIESTHVGVSGM 120
DB 73 TEOCPSPQLTARLIRFWL-----GQKACAYLTEDTDDAVALDHLSTIERAHIVGASMG 128
QY 121 GMIAQILGAKHGRVKSLLTMITSSGNPRMPAPRQVLOKFMVPSKMDKEWIKYNLEL 180
DB 129 GMIAQILFAARFFTRSLAVFFSSNNRPLPPAPRALALLTGPPDTSRRDVVVDNVVR 188
QY 181 LTTI-GSP--GLDRKALDVRKSIERCLCEGTQORQLAAILQSGSRVKLLRRIAVPTLV 237
DB 189 VTKITGSPLYRMPPEQVTRTAAEIDRSFYPLGVSRQFSAILGSGSLHYNQRIIAPTIV 248
QY 238 IESGADPLPYOCGRDIAHPIGARFELIEGMDHPIPERHIPRIELIAGHAAA 292
DB 249 IHGRADKLVPSGSGRAVARAITGARLVLPDGMGHDLPQQLMDQAVGLWSNFAKA 303

RESULT 5
A57139
rnc protein - Streptomyces purpurascens
C/Species: Streptomyces purpurascens
C/Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 09-Jul-2004
C/Accession: A57139
R;Niemi, J.; Maentzela, P.
J. Bacteriol. 177, 2942-2945, 1995
A/Title: Nucleotide sequences and expression of genes from Streptomyces purpurascens t1
A/Reference number: A57139; MUID:95270621; PMID:7751313
A/Accession: A57139
A/Status: preliminary

```

RESULT 7  
T36181

C:Superfamily: carboxyl esterase

RESULT 8  
I60717  
streptothricin-acetyl-transferase - Escherichia coli  
C;Species: Escherichia coli  
C;Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 09-Jul-2004  
C;Accession: I60717  
R;Tietze, E.; Brevet, J.  
Plasmid 25, 217-220, 1991

A;Title: The trimethoprim resistance transposon Tn7 contains a cryptic streptothricin re  
A;Reference number: 160717; MUID:92021239; PMID:1656477  
A;Accession: I60717  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-459 <RES>  
A;Cross-references: UNIPROT:Q47625; UNIPARC:UPI00000AF08C; EMBL:X56815; NID:g499063; PII  
C;Superfamily: streptothricin acetyltransferase

Query Match 17.8%; Score 268.5; DB 2; Length 459;  
Best Local Similarity 29.6%; Pred. No. 1.3e-14;  
Matches 86; Conservative 56; Mismatches 108; Indels 41; Gaps 13;

QY 3 QVKGANGITLVEEEOGHRHPSMLLMGLGQLIDWPPEFIRGLAERGFVFCFNRDAGL 62  
DB 20 KVVVKALSYTESTGDPNHEPILIMG-AMSAVWPDEFCSQLAKMGYVIRYDHR-AGK 77  
QY 63 STKLEGVKKPNIAVFLLASMLKRPVPTLDDMALDVTGLMDALGISTHVVGVSMGGM 122  
DB 78 STSYE-----PQAPYSVEELA-DDVVRVDGYGLEAAHLVGMSLGGF 118  
QY 123 IAQILGAKHGRVKSILMIT---SSGNPRMPAPRPQVQLQKFMKVPKS--MDKEWIKYN 177  
DB 119 LSQLV-ALNPRKRVKSLTIASERLADADPDMPA-FAAIIIEYHQAESLDWSDRDVAVYQ 176  
QY 178 LELTTTGSP-GLDREKALDVRKSIERCLCECTQRLAALLOS--GSRVKLLRI--A 232  
DB 177 VAWRINGTAHAPAEKIONIAELNFD-----TPNLTFTNHTTUGGGERMIGRLNET 229  
QY 233 VPTLVISGAEDPLFPYOCGRDIADHIPCARFELTEGMDHIPERHILPRLIE 283  
DB 230 VPTLIHGTEDPVLPYVHGLAKE-ARGSKMLTEGTGHEHLEDWRIIQ 279

RESULT 9  
A87193  
probable hydrolase [imported] - Mycobacterium leprae  
C;Species: Mycobacterium leprae  
C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C;Accession: A87193  
R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho  
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,  
eam, M.A.; Rutherford, K.M.  
A;Authors: Ruter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S  
Nature 409, 1007-1011, 2001  
A;Title: Massive gene decay in the leprosy bacillus.  
A;Reference number: A86909; MUID:21128732; PMID:11234002  
A;Accession: A87193  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-265 <STO>  
A;Cross-references: UNIPROT:Q9CBBI; UNIPARC:UPI00000C6E55; GB:AL450380; NID:g13093907; F  
C;Genetics:  
A;Gene: ML2269  
C;Superfamily: tropinesterase

Query Match 15.5%; Score 234.5; DB 2; Length 265;  
Best Local Similarity 26.9%; Pred. No. 4.4e-12;  
Matches 83; Conservative 41; Mismatches 104; Indels 81; Gaps 12;

QY 9 ITLVEEEOGHRHPSMLLMGLGQLIDWPPEFIRGLAERGFVFCFNRDAGLSTKLE 68  
DB 2 INLAYEDRGTE--PVVFIAGRGAGRTWQHPQVAPFLAGYRVITFDNRGIGATENTEG 59  
QY 69 VKENIARVFLLASMLKRPVPTLDDMALDVTGLMDALGISTHVVGVSMGGMIAQILG 128  
DB 60 -----PITQTMVADTAVLIESGLGAPVARIIVGVSMGSGFIAQLM 97  
QY 129 AKGGERVKSILMITSSGNPR-----MPA-----PPQVQLQKFMKVPKS 167  
DB 98 VARPELVRAAVLMATRGRLDRTQFFHAAEAEFHDGSLQFPGYNKAVRLLENLSR--KT 155  
QY 168 MDKE----EMI-KYNLELLITIGSPGLDREKALDVRKSIERCLCECTQRLAALLOSQ 222

DB 156 LNDDEVAVADWIANFNNWPIKS--TPGL-----RC-----QTDVA---PQN 190  
QY 223 SRVKLLRRIRIAPVTLVISGAEDPLFPYOCGRDIADHIPCARFELIEGMDH---DIPRHH 278  
DB 191 NRPAYRSIAAPVLYVGFADVVTPSPSLGREVADVLNPGRYLQTPDAGHLGFFERPEAVN 250  
QY 279 PRLIELIAG 287  
DB 251 AAALQFFAG 259

RESULT 10  
C96028  
probable 3-oxoadipate enol-lactonase (EC 3.1.1.24) [imported] - Sinorhizobium meliloti  
C;Species: Sinorhizobium meliloti  
C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
C;Accession: C96028  
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hern  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A;Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing en  
A;Reference number: A95842; MUID:21396508; PMID:11481431  
A;Accession: C96028  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-268 <KUR>  
A;Cross-references: UNIPROT:Q92TMO; UNIPARC:UPI00000CB8A6; GB:AL591985; PIDN:CAC49891.  
A;Experimental source: strain 1021, megaplasmid pSymB  
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, P.; Barloy-Huble  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelauc  
hebault, P.; Vandenberg, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, .  
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A;Reference number: A96039; MUID:21368234; PMID:11474104  
A;Contents: annotation  
C;Genetics:  
A;Gene: pcad; Smb20579  
A;Genome: plasmid  
C;Superfamily: carboxylic ester hydrolase  
C;Keywords: carboxylic ester hydrolase

Query Match 15.2%; Score 229.5; DB 2; Length 268;  
Best Local Similarity 27.2%; Pred. No. 1.2e-11;  
Matches 81; Conservative 43; Mismatches 95; Indels 79; Gaps 10;

QY 1 MAQVKANGITLVEEEOGH-RHHPMSMLLMGLGQLIDWPPEFIRGLAERGFVFCFNRD 59  
DB 1 MQPTRINDVTIHYRVVGVAVTEKPPALVFINSGLTDFRWRDVLRLAGD--FAIVLYDKRG 58  
QY 60 AGLSTKLEGVKKPNIARVFLLASMLKRPVPTLDDMALDVTGLMDALGISTHVVGVSM 119  
DB 59 HGLS-----DYG---QVPYSIEDHATDLAGLLDRLLAVKQIVCGLSV 97  
QY 120 GGMIAQILGAKHGERVKSILMITSSGNPRMPAPRPQVQLQKFMKVPKSMKKEWIKYNLE 179  
DB 98 GGLIAQSLVGRPDVLRVAVLSGTA-----HKIGTVEWDA 133  
QY 180 LTTTIGSPGLDREKALDVRKSIERCLCECTQRLAALLOSRSVKLLR----- 229  
DB 134 RITAIEAHGI--EAVADV---LERWFTFAFRPENLAF--TGYRNNLVKRPQVPGVVGTC 186  
QY 230 -----RIAVPTLVISGAEDPLFPYOCGRDIADHIPCARFELIEGMDHDP 274  
DB 187 AAIRDADFTAAAGRIAPVLCVVGDDGSTPPDPLVRSTADLIPGARFEVIRGAGH-IP 243

RESULT 11  
E70548  
probable bpoC protein - Mycobacterium tuberculosis (strain H37Rv)  
C;Species: Mycobacterium tuberculosis  
C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004

C;Accession: E70548  
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
A;Comor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrrell, B.G.  
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A;Reference number: A70500; MUID:98295987; PMID:9634230  
A;Accession: E70548  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-262 <COL>  
A;Cross-references: UNIPROT:O06420; UNIPARC:UPI000000D5E6A; GB:Z95558; GB:AL123456; NID:9  
A;Experimental source: strain H37Rv  
C;Genetics:  
A;Gene: bpoC  
C;Superfamily: tropinesterase

Query Match 15.1%; Score 228.5; DB 2; Length 262;  
Best Local Similarity 28.1%; Pred. No. 1.4e-11;  
Matches 81; Conservative 29; Mismatches 103; Indels 75; Gaps 10;

Qy 9 ITLEYEQGHRHPSMLLIGLQGLIDWPPEFIRGLAERGFRVFCFDRDAGLSTKLEG 68  
Db 2 INLAYDNG--TGPVVFVIAAGRGAGRTWHPQVPAFLAAGYRCITFDNRGIGATENAEG 59  
Qy 69 VKKPNIARVFLASMGKLPVPTLDDMALDTVGLMDALGIESTHVVGSMGMIAQILG 128  
Db 60 -----FTQTWADTAALIEITLDIAPARVVGVSMGAPFAIAQELM 97  
Qy 129 AKGGERVKSITLMT-----SSGNRMPA---PRPVQLQKFMVPKS 167  
Db 98 VVAPELVSSAVLMATRGRLDRAROFFNKAEBLYDSGVQLPPTVDARALLENFSR--KT 155  
Qy 168 MDKE-----EWKYNLELLTTIGSPGLDREKALDVRKSIERCLCEGTORQLAAILQSGS 223  
Db 156 LNDVAVGDVIAW-FSWWPIKSTPL-----RC-----QLDCAPQT-N 191  
Qy 224 RVKLLRIAVPTLVISGAEDPLLPYQCGRIADHIPGARFELIEGMGH 271  
Db 192 RLPAYRNAAPVLVIGPADDVTPPYLGLREVADALPNRVLQIPDAGH 239

RESULT 12  
T30594  
conserved hypothetical protein PCZA361.30 - Amycolatopsis orientalis  
C;Species: Amycolatopsis orientalis  
C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004  
C;Accession: T30594  
R;Van Wageningen, A.; Kirkpatrick, P.; Williams, D.; Harris, B.; Kershaw, J.; Lennard, N.  
Chem. Biol. 3, 155-162, 1998  
A;Title: Sequencing and analysis of genes involved in the biosynthesis of a vancomycin g  
A;Reference number: Z18804  
A;Accession: T30594  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-276 <VAN>  
A;Cross-references: UNIPROT:O52809; UNIPARC:UPI000000BA4D0; EMBL:AJ223998  
C;Superfamily: carboxyl esterase

Query Match 15.1%; Score 227.5; DB 2; Length 276;  
Best Local Similarity 27.3%; Pred. No. 1.8e-11;  
Matches 84; Conservative 39; Mismatches 130; Indels 55; Gaps 9;

Qy 1 MAQKANGITLEYEQGHRHPSMLLIGLQGLIDWPPEFIRGLAERGFRVFCFDRD 60  
Db 1 MLMTTETGIRLSYHD--HGDGSPVLLTGTGAPSSVWDLHQVPLVRAAGFRVITMNR-- 56  
Qy 61 GLSTKLEGVKKPNIARVFLASMGKLPVPTLDDMALDTVGLMDALGIESTHVVGSMG 120  
Db 57 GIPPSDEGT-----GFTDIDLVDVAALIEHLGVAPCRVVGTSMG 97  
Qy 121 GMIAQILGAKHGERVKSITLMTITSSGNRMPAPRPVQLQKFMVPKSMKEWIKYNLEL 180

C;Accession: E70548  
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
A;Comor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrrell, B.G.  
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A;Reference number: A70500; MUID:98295987; PMID:9634230  
A;Accession: E70548  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-262 <COL>  
A;Cross-references: UNIPROT:O06420; UNIPARC:UPI000000D5E6A; GB:Z95558; GB:AL123456; NID:9  
A;Experimental source: strain H37Rv  
C;Genetics:  
A;Gene: bpoC  
C;Superfamily: tropinesterase

Query Match 15.1%; Score 228.5; DB 2; Length 262;  
Best Local Similarity 28.1%; Pred. No. 1.4e-11;  
Matches 81; Conservative 29; Mismatches 103; Indels 75; Gaps 10;

Qy 9 ITLEYEQGHRHPSMLLIGLQGLIDWPPEFIRGLAERGFRVFCFDRDAGLSTKLEG 68  
Db 2 INLAYDNG--TGPVVFVIAAGRGAGRTWHPQVPAFLAAGYRCITFDNRGIGATENAEG 59  
Qy 69 VKKPNIARVFLASMGKLPVPTLDDMALDTVGLMDALGIESTHVVGSMGMIAQILG 128  
Db 60 -----FTQTWADTAALIEITLDIAPARVVGVSMGAPFAIAQELM 97  
Qy 129 AKGGERVKSITLMT-----SSGNRMPA---PRPVQLQKFMVPKS 167  
Db 98 VVAPELVSSAVLMATRGRLDRAROFFNKAEBLYDSGVQLPPTVDARALLENFSR--KT 155  
Qy 168 MDKE-----EWKYNLELLTTIGSPGLDREKALDVRKSIERCLCEGTORQLAAILQSGS 223  
Db 156 LNDVAVGDVIAW-FSWWPIKSTPL-----RC-----QLDCAPQT-N 191  
Qy 224 RVKLLRIAVPTLVISGAEDPLLPYQCGRIADHIPGARFELIEGMGH 271  
Db 192 RLPAYRNAAPVLVIGPADDVTPPYLGLREVADALPNRVLQIPDAGH 239

RESULT 12  
T30594  
conserved hypothetical protein PCZA361.30 - Amycolatopsis orientalis  
C;Species: Amycolatopsis orientalis  
C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004  
C;Accession: T30594  
R;Van Wageningen, A.; Kirkpatrick, P.; Williams, D.; Harris, B.; Kershaw, J.; Lennard, N.  
Chem. Biol. 3, 155-162, 1998  
A;Title: Sequencing and analysis of genes involved in the biosynthesis of a vancomycin g  
A;Reference number: Z18804  
A;Accession: T30594  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-276 <VAN>  
A;Cross-references: UNIPROT:O52809; UNIPARC:UPI000000BA4D0; EMBL:AJ223998  
C;Superfamily: carboxyl esterase

Query Match 15.1%; Score 227.5; DB 2; Length 276;  
Best Local Similarity 27.3%; Pred. No. 1.8e-11;  
Matches 84; Conservative 39; Mismatches 130; Indels 55; Gaps 9;

Qy 1 MAQKANGITLEYEQGHRHPSMLLIGLQGLIDWPPEFIRGLAERGFRVFCFDRD 60  
Db 1 MLMTTETGIRLSYHD--HGDGSPVLLTGTGAPSSVWDLHQVPLVRAAGFRVITMNR-- 56  
Qy 61 GLSTKLEGVKKPNIARVFLASMGKLPVPTLDDMALDTVGLMDALGIESTHVVGSMG 120  
Db 57 GIPPSDEGT-----GFTDIDLVDVAALIEHLGVAPCRVVGTSMG 97  
Qy 121 GMIAQILGAKHGERVKSITLMTITSSGNRMPAPRPVQLQKFMVPKSMKEWIKYNLEL 180

Db 98 SYIAQELALAHPELLDVAVLMACG-----RSSLVQRVL-----AEGAKLIELTEL 145  
Qy 181 -----LTTIGSPGLDREKLA---LDVRKSIERCLCEGTORQLAAILQSGSRVKL 227  
Db 146 PPGYLAARAHMNLGPATLADDDLTGDLWDLFEASDNWNGPVRAQLQLSAL---PDRIDA 202  
Qy 228 LREIAVPTLVISGAEDPLLPYQCGRIADHIPGARFELIEGMGH---DIPERHIPBLIE 283  
Db 203 YRAIKVPCHVISEFHDLVAPPSAGRELAAAIATHRTIPGCGHFGYLENPEAVNRELLR 262  
Qy 284 LIAGHAAA 291  
Db 263 FIRAESAA 270

RESULT 13  
T35708  
hydrolase - Streptomyces coelicolor  
C;Species: Streptomyces coelicolor  
C;Date: 03-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 09-Jul-2004  
C;Accession: T35708  
R;Murphy, L.; Harris, D.; Parkhill, J.; Barrrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, January 1998  
A;Reference number: Z21548  
A;Accession: T35708  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-261 <MUR>  
A;Cross-references: UNIPROT:O54172; UNIPARC:UPI000000DABAF; EMBL:AL021411; PIDN:CAA1620  
A;Experimental source: strain A3(2)  
C;Genetics:  
A;Gene: SCORDB:SC7H1.13  
C;Superfamily: tropinesterase

Query Match 14.4%; Score 217.5; DB 2; Length 261;  
Best Local Similarity 27.4%; Pred. No. 1.1e-10;  
Matches 69; Conservative 27; Mismatches 67; Indels 89; Gaps 7;

Qy 49 GFRVICFDRDAGLSTKLEGVKKPNIARVFLASMGKLPVPTLDDMALDTVGLMDALG 108  
Db 50 GIRAISVDQDSCIT-----VNPVPVTPPEVLADDLVDLLDALG 88  
Qy 109 IESTHVVGSMGMIAQILGAKHGERVKSITLMTITSSGNRMPAPRPVQLQKFMVPKSM 168  
Db 89 LARAHLLGTSFCGVAQAHAARHPERVASLVLVATTTPSYAMGSAATDELLE-----M 140  
Qy 169 DKEEMIKYNLELLTTIGSPGLDREKALDVRKSIERCLCEGTORQLAAILQSGSRVK-- 226  
Db 141 SHE-----DROQNAADY-----FFTPEG-----QAQQRARPA 167  
Qy 227 --LLRR-----IAVPTLVISGAEDPLLPYQCGRIADHIPGA 261  
Db 168 GYLTRTPGVSRRHDAARRHEIRDGLGGITAPTLVHGTDRLAPYEGALLMERRIPNA 227  
Qy 262 RFELIEGMGHDI 273  
Db 228 ELCPIEGGRHGI 239

RESULT 14  
AG3343  
chloride peroxidase (EC 1.11.1.10) [imported] - Brucella melitensis (strain 16M)  
C;Species: Brucella melitensis  
C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 12-Jul-2004  
C;Accession: AG3343  
R;DelVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova  
Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Lete.  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melite  
A;Reference number: AD3252; PMID:11756688  
A;Accession: AG3343  
A;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-256 <KUR>

A:Cross-references: UNIPROT:Q8VHR3; UNIPARC:UPI0000057DC7; GB:AE008917; PIDN:AAU51914.1;

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEI0733

A:Map position: I

C:Superfamily: tropinesterase

C:Keywords: oxidoreductase

Query Match 14.0%; Score 211.5; DB 2; Length 256;  
Best Local Similarity 28.2%; Pred. No. 3.5e-10;  
Matches 82; Conservative 27; Mismatches 91; Indels 91; Gaps 12;

Qy 7 NGITLEYEQGHRHPSMLLMGL-CGQLIDWPEE-FIRGLAERGFRVFCFDNRDAGLST 64

Db 12 DGLRLAYRQAG--EGDPILLIHGFASSLVNWSVPGWFRFTLTEAGRYVIAIDNRHGFS 69

Qy 65 KLEGVKKPNIA RVLLASMG LKPRVPYTLDDMALDTVGLMDALGIESTHVGVSMGGMIA 124

Db 70 KSHKAE-----DYTPSKMAGDAALDLHLGIAKAVHVGYSMGARIS 110

Qy 125 QILGAKHGERVKS LTL-----MITSSGN-----PRMPA-----PRPQVLQKFMVRP 165

Db 111 AVLAIEHSERVHSAVFGGLGIGMTGADWEPIGEALLAEDPATITHPRGOMFRKPADQT 170

Qy 166 KSMCKEWEIKYNLELLTIGSPGLDREKLALDVKSIERCLCEGTORQLAAILQSGSRV 225

Db 171 RS-----DRIALAACVITSKE--LVP-----V 190

Qy 226 KLARRIAVPTLVISGAEDPLLPYQCGRDIADHIPGARFELIEGMGHDPER 276

Db 191 AAIERILQPVLVAVGTTDDIA--GSAQELANLLP-----NGEALDIPGR 232

# RESULT 15

S34609

carboxylesterase (EC 3.1.1.1) - Pseudomonas sp. (strain KWI-56)

C:Species: Pseudomonas sp.

A:Variety: strain KWI-56

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 12-Jul-2004

C:Accession: S34609

R;Shimada, Y.; Nagao, T.; Sugihara, A.; Iizumi, T.; Yui, T.; Nakamura, K.; Fukase, T.; T

Biochim. Biophys. Acta 1174, 79-82, 1993

A:Title: Cloning and sequence analysis of an esterase gene from Pseudomonas sp. KWI-56.

A:Reference number: S34609; MUID:93326640; PMID:8334166

A:Accession: S34609

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-262 <SHI>

A:Cross-references: UNIPROT:Q9Z3U8; UNIPARC:UPI0000087399; GB:D14529; NID:9397834; PIDN:

C:Superfamily: tropinesterase

C:Keywords: carboxylic ester hydrolase

Query Match 14.0%; Score 211; DB 2; Length 262;  
Best Local Similarity 26.4%; Pred. No. 4e-10;  
Matches 77; Conservative 44; Mismatches 107; Indels 64; Gaps 9;

Qy 7 NG-ITLEYEQGHRHPSMLLMGLGQGLIDWPEEFIRGLAERGFRVFCFDNRDAGLSTK 65

Db 5 NGNVNLSYDVANGHE--CIFFIAGTASDKSNW--DGLRQELSGKYRTVAFDNRDSGESTI 60

Qy 66 LEGVKKPNIA RVLLASMG LKPRVPYTLDDMALDTVGLMDALGIESTHVGVSMGGMIAQ 125

Db 61 CD-----QPYTMDLDAKDALSVMDAEGIQKAAHIVGHSLGGMIAQ 99

Qy 126 ILGAKHGERVKS LTLMITSSGNPRMPAPQVLQKFRVPSKMDKEWIKYNLELLTTIG 185

Db 100 ELAILAPDRVSTLSLVNTAS---RIDNYSVIELARDNSKITTDQRLNRSLYFL---- 152

Qy 186 SPGLDREKLALDVKRSIERCLCEGTORQLAAAILQSGSRVKLLRR----- 230

Db 153 --ALGSKALGSDIFNQVV-----DFASGSQSQPREALIRQWEIDLTVTDTDLSL 200

Qy 231 IAVPTLVISGAEDPLLPYQCGRDIADHIPGARFELIEGMGH-----DIPERHI 278

Db 201 INAKTHVIWASEDKIVTKDQOKMLVNGISGAKFTCIIESGHPMIEAPEEFI 252

Search completed: April 13, 2006, 13:00:03

Job time : 42 secs



Result No.	Query			DB	ID	Description	
	Score	Match	Length				
1	689	45.7	296	2	Q4TR38_9SPHN	Q4tr38 erythroba	
2	623.5	41.3	438	2	Q9KIU0_9BACT	Q9kiu0 uncultured	
3	589.5	39.1	312	2	Q79A78_ACICA	Q79a78 acinetoba	
4	589.5	39.1	312	2	Q57172_ACIDA	Q57172 acinetoba	
5	579	38.4	306	2	Q6D508_ERWCT	Q6d508 erwinia car	
6	562.5	37.3	328	2	Q8HY40_PSRAR	Q8hy40 pseudomonas	
7	562.5	37.3	328	2	Q88NY9_PSPBK	Q88ny9 pseudomonas	
8	561.5	37.2	397	2	Q4K713_PSPBF	Q4k713 pseudomonas	
9	556.5	36.9	299	2	Q97TK4_CLOAB	Q97tk4 clostridium	
10	552.5	36.6	304	2	Q89R88_BRAJA	Q89r88 bradyrhizob	
11	548.5	36.3	330	2	Q87XK0_PSESM	Q87xk0 pseudomonas	
12	542	35.9	345	2	Q4FSG6_PSGAM	Q4fsg6 pseudomonas	
13	540.5	35.8	330	2	Q4ZPH7_PSPSY	Q4zph7 pseudomonas	
14	536.5	35.6	328	2	Q4J684_AZОВI	Q4j684 azotobacter	
15	488	32.3	312	2	Q59ZC7_9TRYP	Q59zc7 trypanosoma	
16	485	32.1	301	2	Q96935_MYCTU	Q96935 mycobacteri	
17	485	32.1	301	2	Q7ULJ8_MYCBO	Q7ulj8 mycobacteri	
18	485	32.1	336	2	Q7D9H6_MYCTU	Q7d9h6 mycobacteri	
19	475.5	31.5	304	2	Q9CBK5_MYCLE	Q9cbk5 mycobacteri	
20	474.5	31.4	302	2	Q73SF6_MYCPA	Q73sf6 mycobacteri	
21	473.5	31.4	291	2	Q88J45_PSPBK	Q88j45 pseudomonas	
22	441	29.2	311	2	Q5LV17_SILPO	Q5lv17 silicibacte	
23	426.5	28.3	298	2	Q5MD37_9DELT	Q5md37 cyatobacter	
24	425.5	28.2	291	2	Q9APW4_PSPAR	Q9apw4 pseudomonas	
25	425	28.2	346	2	Q4Q8A8_LEIMA	Q4q8a8 leishmania	
26	378	25.0	716	2	Q4QE85_LEIMA	Q4qe85 leishmania	
27	376	24.9	280	2	Q7SQP7_ECOLI	Q7sqp7 escherichia	
28	375.5	24.9	432	2	Q4QE87_LEIMA	Q4qe87 leishmania	
29	372	24.7	280	2	Q5Q6Z8_ECOLI	Q5q6z8 escherichia	
30	372	24.7	280	2	Q7SQQ0_SALEN	Q7sqq0 salmonella	
31	370	24.5	280	2	Q75WM3_ECOLI	Q75wm3 escherichia	

```

ID O9KIUO_9BACT PRELIMINARY; PRT; 438 AA.
AC Q9KIUO;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Esterase.
OS unclutered bacterium.
OC Plasmid pH117.
OC Bacteria; environmental samples.
OX NCBI_TaxID=77133;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20336470; PubMed=10877816;
RX DOI=10.1128/EM.66.7.3113-3116.2000;
RA Henne A., Schmitz R.A., Bomeke M., Gottschalk G., Daniel R.;
RT "Screening of environmental DNA libraries for the presence of genes
RT conferring lipolytic activity on Escherichia coli.";
RL Appl. Environ. Microbiol. 66:3113-3116(2000).
DR EMBL: AF223648; AAF87665.1; -; Genomic DNA.
DR GO: GO:0016787; F:hydrolase activity; IEA.
DR InterPro: IPR000073; A/b hydrolase.
DR Bacteria; environmental samples.
DR InterPro: IPR000379; Ser_estrs.
DR Pfam: PF00561; Abhydrolase_1; 1.
KW Hydrolase; Plasmid.
SQ SEQUENCE 438 AA; 46539 MW; DA36E26AF585B630 CRC64;

Query Match 41.3%; Score 623.5; DB 2; Length 438;
Best Local Similarity 47.2%; Pred. No. 3.2e-39;
Matches 141; Conservative 39; Mismatches 114; Indels 5; Gaps 3;

QY 1 MAOVKANGITLVEEQGHRHPSMLLGLGGLIDWPPEEFGIRLAERGFRVLCFNDA 60
DB 132 MPQAOAGLFDYEDFGPPAPCVLLIMGLGMPAALWPDFAVDGLRAARLRVIRFNDC 191

QY 61 GLSTKLGKGVKPNIRVFLASMLGKRPVPTLDDMALDTVGLMDALGIESTHVGVSMG 120
DB 192 GHSSKLGCTPTLLPAIARALLLPVAPVPTLDDMADDTAALLAAGVERAHVVGASMG 251

QY 121 GMTAIIAGKHGKRVKSLTMITSSGN--PRMPAPRPQVLOKFMVRPKSMDEKWKYNL 178
DB 252 GMTAIVLAARHPDRLVSLTSIMSSGNPSRIALGSRRLRALILHRPSRTDDVDALTAHL 311

QY 179 -ELLTTIGSPGL--DREKALDVRKSIERCLCEGTQORQLAAILQSGSRVKLRLRAVPT 235
DB 312 VQVFGVIGSPGSDHQALHQQLRVVQAGVTHPAGTARQLAAILAGDRKKLARLEVP 371

QY 236 LVISGADPLLPYQCGRDIAHPIGAFELIEGMDHPIERHPIRLIELIAGHAAAEA 294
DB 372 LVTHGGDDPLVPLAAGIDTAKHIGARLKVIPGMDHDFPALQPELAGLISAHIASTA 430

RESULT 3
Q79A78 ACICA PRELIMINARY; PRT; 312 AA.
AC Q79A78;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Carboxyl esterase.
OS Name=estB;
OS Acinetobacter calcoaceticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=471;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=BD413;
RA Kok R.G., Bart A., Hellingwerf K.J.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: X88895; CAA61351.1; -; Genomic DNA.
DR GO: GO:0016787; F:hydrolase activity; IEA.
DR InterPro: IPR000073; A/b hydrolase.
DR InterPro: IPR000379; Ser_estrs.

```

```

DR Pfam: PF00561; Abhydrolase_1; 1.
KW Hydrolase.
SQ SEQUENCE 312 AA; 34681 MW; 14E020BA7654D905 CRC64;

Query Match 39.1%; Score 589.5; DB 2; Length 312;
Best Local Similarity 45.5%; Pred. No. 8.6e-37;
Matches 130; Conservative 53; Mismatches 100; Indels 3; Gaps 3;

QY 6 ANGITLEYEEQGHRHPSMLLGLGGLIDWPPEEFGIRLAERGFRVLCFNDAJLSTK 65
DB 24 SNGLELHVEVGNDPHTILLIMGLGGLIDWPPEEFGIRLAERGFRVLCFNDAJLSTK 83

QY 66 LEGV-KKPNIRVFLASMLG-KRPVPTLDDMALDTVGLMDALGIESTHVGVSMGMI 123
DB 84 IRHKGRLTLKMSRFTILGNGQAPYLYDAEDVSLLEMRKKNVNVIGASMGMI 143

QY 124 AQILGKHGKRVKSLTMITSSGNPRMPAPRPQVLOKFMVRPKSMDEKWKYNLELLTT 183
DB 144 AQIIAAKYPEKVEKLMFTSNQPLPPFPKQFLSLIGKPKSDEDEGIINHSLKLF 203

QY 184 ICSPLGDLREKALD-VKRSIERCLCEGTQORQLAAILQSGSRVKLRLRAVPTLVISGAE 242
DB 204 ICSPGYVNVQVEAIQTARKLYQRSYHPAGVLOQFLAILCTGSLQDLQKISQPTLVINGSR 263

QY 243 DPLLPYQCGRDIAHPIGAFELIEGMDHPIERHPIRLIELIAGH 288
DB 264 DRLLPSSHGKAVAKAIGAKFELIQMGHDIHPHFIPQLSYLFAHH 309

RESULT 4
Q57172 ACIAD PRELIMINARY; PRT; 312 AA.
AC Q57172;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE EstB protein (Putative hydrolase).
OS Name=estB; OrderedLocNames=ACIAD1064;
OS Acinetobacter sp. (strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=ADP1;
RX MEDLINE=95400495; PubMed=7670642;
RA Geissdoerfer W., Froesch C.S., Haspel G., Ehart S., Hillen W.;
RT "Two genes encoding proteins with similarities to rubredoxin and
RT rubredoxin reductase are required for conversion of dodecane to lauric
RT acid in Acinetobacter calcoaceticus ADP1.";
RT Microbiology 141:1425-1432(1995).
RN (2)
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=ADP1;
RX MEDLINE=97228433; PubMed=9074511; DOI=10.1016/S0378-1119(96)00728-7;
RA Geissdoerfer W., Ratajczak A., Hillen W.;
RT "Nucleotide sequence of a putative periplasmic Mn superoxide dismutase
RT from Acinetobacter calcoaceticus ADP1.";
RL Gene 186:305-308(1997).
RN (3)
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=ADP1;
RX MEDLINE=98162050; PubMed=9501429;
RA Geissdoerfer W., Ratajczak A., Hillen W.;
RT "Transcription of ppk from Acinetobacter sp. strain ADP1, encoding a
RT putative polyphosphate kinase, is induced by phosphate starvation.";
RL Appl. Environ. Microbiol. 64:896-901(1998).
RN (4)
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=ADP1;
RA Geissdoerfer W., Kok R.G., Ratajczak A., Hellingwerf K.J., Hillen W.;
RT "The genes rubA and rubB for alkane degradation in Acinetobacter sp.
RT strain ADP1 are in an operon with estB, encoding an esterase, and

```



```

QY 4 VKANGITLEYEQGHRHPSMLLIMGLGQLIDWPEEFIRGLAERGRVFCFDRDAGLS 63
Db 31 VELGDVRLAYISIGRSDSPALLVWGLGGQLIHWPDEVVSAACGQGRVIRYDNRDVGLS 90
QY 64 TKLEGVKKPNIARVFLLASMGKPRVPYTLDDMALDVTGLMDALGIESTHVGVSMGMI 123
Db 91 ANWVPVSSRLTYEVVRYLGLPVSAPTYTLTDMAGDALHLLDALDIPQAVLGSAGMGI 150
QY 124 AQILGAKHGERVKSITLMTSSGNPRMPAPQVQLQKFMVPKSMDEWIKYNLELTT 183
Db 151 AQHIADMAPORLLSLTYMTSSGAEGLPAPSESLRLAR-REASREQAEEQQADLLAA 209
QY 184 IGSPL--DREKALDVRKSIERCPCGEGTORQALAAIILQSGSRVKLLRRIAVPTLVISGA 241
Db 210 LGSFEVDRDQQLLQARSDRAFNPEGVRQQLAILAEFSRVPLNRLQVPTLVHGT 269
QY 242 EDPLPYQCGRDIAHIFGARFELIEGMGHDIPIPERHIPRLIELIAGHAAA 291
Db 270 ADPLPVHGVHVAHIRGSELKIPGLAHRFQAEKPELLAAVVPYLKA 319

RESULT 7
Q88N9_PSEPK
ID Q88N9_PSEPK PRELIMINARY; PRT; 378 AA.
AC Q88N9_PSEPK
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DE Hydrolyase, alpha/beta fold family.
GN OrderedLocustNames=PP1064;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22423060; PubMed=12534463;
RX DOI=10.1046/j.1462-2920.2002.00366.x;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapfel E.K., Scanlan D., Tran K.,
RA Moazzes A., Utterback T.R., Rizzo M., Lee K., Kosack D., Meestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AB016778; AAN66689.1; -; Genomic_DNA.
DR TIGR; PP1064; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR InterPro; IPR000073; A/b_hydrolase.
DR InterPro; IPR000379; Ser_gstrs.
DR Pfam; PF00561; Abhydrolase_1; I.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 378 AA; 41119 MW; 055BAB0F98F3EBBA CRC64;

Query Match 37.3%; Score 562.5; DB 2; Length 378;
Best Local Similarity 44.4%; Pred. No. 1.3e-34;
Matches 122; Conservative 47; Mismatches 103; Indels 3; Gaps 2;

QY 3 QVANGITLEYEQGHRHPSMLLIMGLGQLIDWPEEFIRGLAERGRVFCFDRDAGL 62
Db 78 RAELGAVSLVYQSVGAPDPAALLVWGLGGQLIHWPDDVEALCRQGRVIRYDNRDVG 137
QY 63 STKLEGVKKPNIARVFLLASMGKPRVPYTLDDMALDVTGLMDALGIESTHVGVSMGMI 122
Db 138 SRWQVPPHANLTLELLRYLGLPVSAPTYTLTDMAGDALGIHMDALGVQRPHVLSVGMGM 197
QY 123 IAQILGAKHGERVKSITLMTSSGNPRMPAPQVQLQKFMVPKSMDEWIKYNLELTT 182

```

```

Db 198 IAQHLAAMAFVRSITLVWSSGAGLPAPDPAVVOLLAR-RSAPPREVAIEQQADLLA 256
QY 183 TIGSPGL--DREKALDVRKSIERCPCGEGTORQALAAIILQSGSRVKLLRRIAVPTLVISG 240
Db 257 ALGSEPRVDDREVLLHQAAQAYDRAFNPDGAKQIMAILAEPKSRVELLNQLRVPTLVVHG 316
QY 241 AEDPILPYQCGRDIAHIFGARFELIEGMGHDIPI 275
Db 317 TADPLPVHGVHVAHIRGSELKIPGLAHRFQAE 351

RESULT 8
Q4K713_PSEFS
ID Q4K713_PSEFS PRELIMINARY; PRT; 397 AA.
AC Q4K713_PSEFS
DT 13-SEP-2005 (Tremblrel. 31, Created)
DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)
DE Hydrolyase, alpha/beta fold family.
GN ORFNames=PFL4889;
OS Pseudomonas fluorescens (strain Pf-5).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=220664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PF-5;
RX PubMed=15980861; DOI=10.1038/nbt1110;
RA Paulsen I.T., Press C., Ravel J., Kobayashi D., Myers G.S.,
RA Mavrodin D., DeBoy R.T., Seshadri R., Ren Q., Madupu R., Dodson R.J.,
RA Durkin S., Brinkac L.M., Daugherty S.C., Sullivan S.A., Rosovitz M.,
RA Gwinn M.B., Zhou L., Nelson W.C., Weidman J., Watkins K., Tran K.,
RA Khouri H.M., Pierson E., Pierson L. III, Thomasow L., Loper J.;
RT "Complete genome sequence of the plant commensal Pseudomonas
RT fluorescens Pf-5.";
RL Nat. Biotechnol. 23:873-878(2005).
DR EMBL; CP000076; AAY94119.1; -; Genomic_DNA.
KW Hydrolase.
SQ SEQUENCE 397 AA; 42869 MW; 2F77FD2116F6E86 CRC64;

Query Match 37.2%; Score 561.5; DB 2; Length 397;
Best Local Similarity 45.4%; Pred. No. 1.6e-34;
Matches 122; Conservative 43; Mismatches 101; Indels 3; Gaps 2;

QY 9 ITLEYEQGHRHPSMLLIMGLGQLIDWPEEFIRGLAERGRVFCFDRDAGLSTKLEG 68
Db 94 VSLAYQSIGRTSDPALLVWGLGGQLIHWPDEVVVALCOQGRVIRYDNRDVGSLTRQA 153
QY 69 VKKPNIARVFLLASMGKPRVPYTLDDMALDVTGLMDALGIESTHVGVSMGMIQILG 128
Db 154 PVSANLTTFEVLRYKGLPVAAPTYTLTDMAGDALGIHMDALHIEQFHVIGSMGMIAQHLA 213
QY 129 AKHGERVKSITLMTSSGNPRMPAPQVQLQKFMVPKSMDEWIKYNLELTTIGSPG 188
Db 214 AMEPRVESLTLMTSSGAEGLPAPNAALVQLLSR-RNAPSREAAEQQADLLAALGSPK 272
QY 189 L--DREKALDVRKSIERCPCGEGTORQALAAIILQSGSRVKLLRRIAVPTLVISGADPLL 246
Db 273 VSDRQVLLQAAVAYDRAFNPEGVRQQLAILAEFSRVALLKQLRVPTLVHGTADPLL 332
QY 247 PYQCGRDIAHIFGARFELIEGMGHDIPI 275
Db 333 PVMHGVHVAHIRGSELKIPGLAHRFQAE 361

RESULT 9
Q97TK4_CLOAB
ID Q97TK4_CLOAB PRELIMINARY; PRT; 299 AA.
AC Q97TK4_CLOAB
DT 01-OCT-2001 (Tremblrel. 18, Created)
DT 01-OCT-2001 (Tremblrel. 18, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)

```

```
DE Carboxyl esterase, a/b hydrolase.
GN OrderedLocusNames=CAP0097;
QS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 824 / DSM 792 / VRM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA DOI=10.1128/JB.183.16.4823-4838.2001;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R.L., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L.A., Soucaille P.,
RA Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.;
RA "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum."
RL J. Bacteriol. 183:4823-4838(2001).
RL EMBL; AB001438; AAK76842.1; -; Genomic DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:006725; P:aromatic compound metabolism; IEA.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR003089; AB_hydrolase.
DR Pfam; PF00561; Abhydrolase_1; 1.
DR PRINTS; PR00111; ABHYDROLASE.
KW Complete proteome; Hydrolase; Plasmid.
SQ SEQUENCE 299 AA; 33303 MW; DD958146552C420E CRC64;

Query Match 36.9%; Score 556.5; DB 2; Length 299;
Best Local Similarity 41.1%; Pred. No. 2.7e-34;
Matches 118; Conservative 53; Mismatches 113; Indels 3; Gaps 2;

Qy 1 MAQVKGITLEYEEQGRHHPHSMGLGQQLIDWPEEIRGLAERGRVFCFNRDA 60
Db 1 MAHVNTGQIQEYVFGRTNPTVLVAGNQAQNFWSDFCEMLAQNQVIRFNRDA 60

Qy 61 GLSTKLGKVPKPNRIARVFLASMLGKPRVPYTLDDMALDTVGLMDALGIESTHVGVSMG 120
Db 61 GLSTKFDAGIPDMSKIYQAQEQKPIKTATYLEDMAADDVAGLLDALEIKKACGASMG 120

Qy 121 GMIAQILGAKHGERVKSILMITSSGNRMPAPRPQVLOKFMVVKSMOKEEWIKYNLEL 180
Db 121 GMIAQVAYRPLRICLSLISMSGTGNPNPKISQETLEIVTATPPN-QRGAVIETVHM 179

Qy 181 LTTIGSPG--LDREKALDVKSTIERCLCEPGTORQALAILQSGRVKLLRRVAVPTLVI 238
Db 180 WKLLWSKGFPEESRAIRYTESYDRCPYQGAVRQNAALVANGDRKRLSLLRVPTLVI 239

Qy 239 SGAEDPLLPVOCGRDIADHIPGARFELIEGMDHDIPERHIPRLIELIAGHA--AAEA 294
Db 240 HGTADPLIPVEAGKDTARTIPNAKLLIEGMDHMPKGTWCIVEAI 286

RESULT 10
Q89R88 BRAJA PRELIMINARY; PRT; 304 AA.
ID Q89R88 BRAJA
AC Q89R88;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE B112894 protein.
GN OrderedLocusNames=b112894;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,

RA Saemoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpō S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RL DNA Res. 9:189-197(2002).
DR EMBL; BA000440; BAC48149.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00561; Abhydrolase_1; 1.
KW Complete proteome.
SQ SEQUENCE 304 AA; 33095 MW; 58552CC0F028B6A8 CRC64;

Query Match 36.6%; Score 552.5; DB 2; Length 304;
Best Local Similarity 44.6%; Pred. No. 5.6e-34;
Matches 133; Conservative 42; Mismatches 112; Indels 11; Gaps 4;

Qy 4 VKANGITLEYEEQGRHHPHSMGLGQQLIDWPEEIRGLAERGRVFCFNRDAGLS 63
Db 10 VRANGIDICVEIFGNDNAEPLLLIMGLGAQMIHWDADFCEQLAVHGPRVIRFNRDICKS 69

Qy 64 TKLEGVKKPNRIARVFLASMLGKPRVPYTLDDMALDTVGLMDALGIESTHVGVSMGMI 123
Db 70 SHLTGGRRLTPELLKRLRIPVAATYKLIWDARTVGLMDALGKSAHLVGASMGMI 129

Qy 124 AQILGAKHGERVKSILMITSSGNRMPAPRPQVLOKFMVVKSMOKEEWIKY--NLELL 181
Db 130 AQEVTLSPFPRVRSLSIMSTTGNRPPPTTREAAMLMAPPKRSKEEFVRYQTWKVL 189

Qy 182 TTTGSGLDREKALD---VRKSIERCLCEPGTORQALAILQSGRVKLLRRVAVPTLVI 238
Db 190 RAGAFP---EEBALPDRAERVFARGLNPAGVGRQLRAVLASGSRKERLHAMKTPTLVI 245

Qy 239 SGAEDPLLPVOCGRDIADHIPGARFELIEGMDHDIPERHIPRLIELIAGHA--AAEA 294
Db 246 HGTVDPLVHVEGKGKDTAASIPGAKLLMVGEONGHALPWFPEIIGADKARRAAQA 303

RESULT 11
Q87XKO PSFSM
ID Q87XKO PSFSM PRELIMINARY; PRT; 330 AA.
AC Q87XKO;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hydrolase, alpha/beta fold family.
GN OrderedLocusNames=PSPT04178;
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OX NCBI_TaxID=323;
RN [1]
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=DC3000;
RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;
RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA Gwinn M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,
RA Nelson W.C., Daviden T.M., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H.M., Fedorova N.B., Tran B., Ruseell D., Berry K.J.,
RA Uterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,
RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,
RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,
RA Bender C.L., White O., Fraser C.M., Collier A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
RL Pseudomonas syringae pv. tomato DC3000."
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
DR EMBL; AB016853; AAO57634.1; -; Genomic DNA.
DR TIGR; PSPT04178; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000379; Ser_estrs.
```

[illegible]



```

QY 239 SGAEDPLLPYQCGRDIAADHIPGARFELIEGMGHDIPE 275
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 267 HGTADPLLPVHGWHVAHITQSGQLRLPGLAHRFQB 303
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 14
Q4J684 AZOVI PRELIMINARY; PRT; 328 AA.
AC Q4J684;
DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE Alpha/beta hydrolase fold precursor.
GN ORFNames=AvindRAFT 8586;
OS Azotobacter vinelandii AvOP.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=322710;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AvOP.
RG US DOE Joint Genome Institute (JGI-PGP);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Pliuck S., Richardson P.;
RA "Sequencing of the draft genome assembly of Azotobacter vinelandii
RT AvOP";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AvOP.
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RA "Annotation of the draft genome assembly of Azotobacter vinelandii
RT AvOP";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AvOP.
RG US DOE Joint Genome Institute;
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Pliuck S., Richardson P.;
RA Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAU03000001; EAM06907.1; -; Genomic_DNA.
KW Hydrolase; Signal.
FT SIGNAL
SQ SEQUENCE 328 AA; 36068 MW; 3258374E36CC4A7D CRC64;

Query Match 35.6%; Score 536.5; DB 2; Length 328;
Best Local Similarity 43.9%; Pred. No. 1e-32;
Matches 118; Conservative 44; Mismatches 104; Indels 3; Gaps 2;

QY 9 ITLEYEQGHRHPSMLLMGLGGQLIDWPEEFIRGLAERGFVFCFDRDAGLSTKLEG 68
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 36 VQLAYOSIGRETPALVNLVGLGGQLIDWPEVLGLCRGFRVIFDRDVLGSAWRQE 95
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 69 VKKPNIAVFLASMLGKPRVPTLDDMALDTVGLMDALGIESTHVGVSMGGMIAQIIG 128
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 96 VPKVSLTYGVLRLHGLPVSAPYGLRDMARDITLALMARLQVPRFRFVLGVSMGGMIAQHWA 155
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 129 AKHGERVKSILMITSSGNPRMPAPRQVLOKFMVRPKSMDEKWKYNLELLTTIGSP- 187
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 156 DLAPERVRSLLTMTSSSALGLPGPSAALL-KLLAQREAGSREVAIERQVELLAUSPQ 214
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 188 -GLDREKALDVRKSIERCLCPBGTQRLAAALIQSGSRVKLLRRIAVPTLVISGAEDPL 246
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

215 VEVDREQLRRQAAATAYDRAFNPEGVKRQQLAVLAEPSEVRELLRLPTLVHGTADPLL 274
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 247 PYQCGRDIAADHIPGARFELIEGMGHDIPE 275
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 275 PVMHGWHVAHITQSGQLRLPGLAHRFQB 303
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 15
Q582C7 9TRYP PRELIMINARY; PRT; 312 AA.
AC Q582C7;
DT 10-MAY-2005 (T-EMBLrel. 30, Created)
DT 10-MAY-2005 (T-EMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (T-EMBLrel. 30, Last annotation update)
DE Hydrolase, alpha/beta fold family, putative.
GN ORFNames=TB927.5.2380;
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA Ghedin E., Blandin G., Bartholomeu D., Caler E., Haas B., Hannick L.,
RA Shallom J., Hou L., Djikeng A., Feldblyum T., Hostetler J.,
RA Johnson J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,
RA Khalak H.G., Salizberg S., Simpson A.J., Tallon L., Van Aken S.,
RA Wanless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA El-Sayed N.M., Khalak H., Adams M.D.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA Haas B., Blandin G., El-Sayed N.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC091655; AAX80442.1; -; Genomic_DNA.
KW Hydrolase.
SQ SEQUENCE 312 AA; 34433 MW; ECH85ECF9E41DE87 CRC64;

Query Match 32.3%; Score 489; DB 2; Length 312;
Best Local Similarity 39.0%; Pred. No. 4.9e-29;
Matches 110; Conservative 54; Mismatches 114; Indels 4; Gaps 3;

QY 9 ITLEYEQGHRHPSMLLMGLGGQLIDWPEEFIRGLAERGFVFCFDRDAGLSTKLEG 68
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 21 ISLSYDTFGNSKDPCLLVGLAGVGRVWRDAFCEMIAGKGFVVRVDRDVLGSLTHLN 80
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 69 VKKPNIAVFLASMLGKPRVPTLDDMALDTVGLMDALGIESTHVGVSMGGMIAQIIG 128
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 81 QTPNVNQCLLPQFLSFLRKVPTLIEDMAADGNLLTALGIERAHVGVSMGGMIAQIMA 140
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 129 AKHGERVKSILMITSSGNPRMPAPRQVLOKFMVRPKSMDEKWKYNLELLTTIGSPG 188
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 141 IKYPSRVSRLGIYSHGTSKSRVPETFTSKLLFMKPKSKSALEDVVDVDFKALARHFRPG 200
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 189 --LDREKALDVRKSIERCL-CPBGTQRLAAALIQSGSRVKLLRRIAVPTLVISGAEDPL 245
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 201 YNVDEEFPKLAQEQLRANDYPQGMRLQALAILSAKSEKCLKTITPTLIHGMLEDEV 260
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 246 LPYQCGRDIAADHIPGARFELIEGMGHDIPE 275
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 261 VPYQNGLOIAEAVGPAKLVITYPRMGHEIPVLMFMSISOEIA 302
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Search completed: April 13, 2006, 12:59:18
Job time : 234 secs

```



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 13, 2006, 12:59:33 ; Search time 47 Seconds  
(without alignments)  
517.163 Million cell updates/sec

Title: US-10-686-490D-2  
Perfect score: 1509  
Sequence: 1 MAQVANGTLEYEQGHRH.....ERHPLIELIAGHAAABA 294

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/6\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/H\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/PCUTUS\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/RE\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	607.5	40.3	322	2	US-09-328-352-5739
2	562.5	37.3	422	2	US-09-252-991A-21999
3	483	32.0	360	2	US-09-540-236-3609
4	314.5	20.8	267	2	US-08-935-263-10
5	314.5	20.8	267	2	US-09-594-185-10
6	314.5	20.8	267	2	US-10-033-078-10
7	314.5	20.8	267	2	US-10-763-933-10
8	262.5	17.4	247	2	US-09-902-540-14883
9	192	12.7	287	2	US-09-252-991A-22466
10	189.5	12.6	251	2	US-09-345-469-4
11	189	12.5	268	2	US-09-252-991A-32242
12	185	12.3	265	2	US-09-902-540-11639
13	183	12.1	289	2	US-09-469-211A-6
14	183	12.1	289	2	US-09-446-681-6
15	182.5	12.1	271	2	US-09-902-540-12298
16	181.5	12.0	264	2	US-09-134-001C-5592
17	180	11.9	338	2	US-09-252-991A-32574
18	175	11.6	261	2	US-09-902-540-13754
19	174	11.5	334	2	US-09-902-540-11426
20	174	11.5	786	2	US-09-252-991A-32352
21	173.5	11.5	271	2	US-09-328-352-7546
22	168	11.1	281	2	US-09-902-540-11020
23	163	10.8	287	2	US-09-902-540-15478
24	162.5	10.8	275	2	US-09-902-540-15896
25	157	10.4	320	2	US-10-272-490-6
26	156.5	10.4	274	2	US-09-252-991A-21590
27	155.5	10.3	286	2	US-09-355-166-9

28	155.5	10.3	388	1	US-08-232-519-2	Sequence 2, Appli
29	155.5	10.3	388	1	US-08-456-956-2	Sequence 2, Appli
30	155	10.3	320	2	US-09-252-991A-18301	Sequence 18301, A
31	150	9.9	327	2	US-10-272-490-58	Sequence 58, Appl
32	149.5	9.9	274	2	US-09-949-016-6477	Sequence 6477, Ap
33	149.5	9.9	279	2	US-09-949-016-7659	Sequence 7659, Ap
34	149	9.9	326	2	US-10-272-490-20	Sequence 20, Appl
35	147	9.7	321	2	US-09-252-991A-29844	Sequence 29844, A
36	146	9.7	269	2	US-10-272-490-22	Sequence 22, Appl
37	146	9.7	462	2	US-09-712-363-209	Sequence 209, App
38	144.5	9.6	331	2	US-10-272-490-50	Sequence 50, Appl
39	144.5	9.6	339	2	US-10-272-490-60	Sequence 60, Appl
40	144	9.5	276	2	US-09-489-039A-13894	Sequence 13894, A
41	143	9.5	296	2	US-09-489-039A-17531	Sequence 7531, Ap
42	141.5	9.4	321	2	US-08-909-125-5	Sequence 5, Appli
43	138.5	9.2	312	2	US-09-424-349A-5	Sequence 2, Appli
44	138.5	9.2	554	1	US-08-106-761-2	Sequence 2, Appli
45	138.5	9.2	554	2	US-08-909-125-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1  
US-09-328-352-5739  
; Sequence 5739, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC999-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 5739  
; LENGTH: 322  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-5739

Query Match	40.3%	Score	607.5	DB 2	Length	322
Best Local Similarity	46.9%	Pred. No.	3.1e-57			
Matches	134	Conservative	49	Mismatches	98	Indels 5; Gaps 4;
QY	7	NGITLEYEEQGRHPSMLLIMGLGQGLDWPPEEPIRGLAERGFRVFCFDRDAGLSTKL	66			
DB	33	NGIELHVEVGKPEHPTILLINGLCAQMLFWDFFCKSLIDQGFVIRFDNRDGLSSKV	92			
QY	67	--EGVKPNIAKRVLLASGMLKPR-VPYTLDMDALDTVGLMDALGIESTHVGVSMGMI	123			
DB	93	RHOG-KRLNTWKLMGRFALGRNQAPYTYDMADDSVMLLDRLGVSKARVIGASMGMI	151			
QY	124	AQILGAKGHRVKSJTLMTSSGNRMPAPRQVQLQKMRVPKMDKEWIKYNLELLTT	183			
DB	152	AQILAAYPEKVEKLGMLFTSNQPLPPPPKQLLSLIGKPEDRDEGIVNHSKLFLQ	211			
QY	184	IGSPG-LDREKIALDVRSIERCLCEGTQRLAAAILQSGSRVKLRRIAVPTLVISGAE	242			
DB	212	IGSPGINHIEAVQTKLYQRSYYPAGVQLQFLLCTGSLQLDREIKQPTLVLHGR	271			
QY	243	DPLLPYQGRDIADHIPGARFELIEGMDHIDPERHIPRIELIAGH	288			
DB	272	DRLLPPSHGKAVAKAISGAKFELIDGMGHIDIPAHFIPQLSGLFAHH	317			

RESULT 2  
US-09-252-991A-21999  
; Sequence 21999, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 21999  
 ; LENGTH: 422  
 ; TYPE: PRP  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-21999

Query Match 37.3%; Score 562.5; DB 2; Length 422;  
 Best Local Similarity 43.1%; Pred. No. 3.5e-52;  
 Matches 125; Conservative 45; Mismatches 117; Indels 3; Gaps 2;  
 ;  
 QY 4 VKANGITLVEEQQHRRHPSMLLMGLGGQIDWPPEFIRGLAERGRVFCFNRDAGLS 63  
 DB 125 VELGDVRLAYQSGRSDPALLVMGLGGQIDHPDEVVSALCEQGRVIRYDNRDVGLS 184  
 QY 64 TKLEGVKKPNIRARVFLASMLKPRVPTLDDMALDTVGLMDALGISTHVVGVSMGGM 123  
 DB 185 ANVPVPSRLTYEVVRYRLGLPVSAFYTLTDAGDALHLLDALDIPQAHVGLASMGGM 244  
 QY 124 AQILGAKHGRVSKLTMTTSSGNPRMPAPRQVLOKFMKVPKSMKKEWIKYNLELLT 183  
 DB 245 AQIADMAPQLLSLTVMTSSGAEGLPAPSESILLRLAR-REAAAREQAVEQQAIDLAA 303  
 QY 184 IGSPLG--DREKALDVRKSIERCLCEGTQORQALAAILOSQSRVGLLRRIAVPTLVISGA 241  
 DB 304 LGSFEVDRDQQLLQARSYDRAFNEFEGVQRQLLALAEPSVPLNRLQVPIVHGT 363  
 QY 242 EDPLLPYQCGRDIAHPIPGARFELIEGMGHDIPERHPIRLIELIAGHAAA 291  
 DB 364 ADPLLPVHGVHVAHIRGSELKILPGLAHRFQEPKPELLAAVVPYKA 413

RESULT 3  
 US-09-540-236-3609  
 ; Sequence 3609, Application US/09540236  
 ; Patent No. 6673910  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary L. Breton et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR  
 ; FILE REFERENCE: 2709.2005-001  
 ; CURRENT APPLICATION NUMBER: US/09/540,236  
 ; CURRENT FILING DATE: 2000-04-04  
 ; NUMBER OF SEQ ID NOS: 3840  
 ; SEQ ID NO 3609  
 ; LENGTH: 360  
 ; TYPE: PRP  
 ; ORGANISM: M.catarrhalis  
 ; FEATURE:  
 ; NAME/KEY: UNSURE  
 ; LOCATION: (250)  
 ; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unkno  
 US-09-540-236-3609

Query Match 32.0%; Score 483; DB 2; Length 360;  
 Best Local Similarity 37.3%; Pred. No. 1.1e-43;  
 Matches 107; Conservative 61; Mismatches 113; Indels 6; Gaps 3;  
 ;  
 QY 8 GITLEVEEQHRRHPSMLLMGLGGQIDWPPEFIRGLAERGRVFCFNRDAGLSTK-- 65  
 DB 42 GIKLCVAGNPEHPMIFITLGSQMMFWSDFKRFIDAGFFVIRFNRDGLSKTIQ 101  
 QY 66 LEGVKKPNIRARVFLASMLKPR---VPTLDDMALDTVGLMDALGISTHVVGVSMGGM 122  
 DB 102 IDGLPRLNTPKMLKMQAGLSNRSEPVAYTLTDMAEDAAARLKTMLHNVNLIGASMGGM 161

QY 123 IAQILGAKHGRVSKLTMTTSSGNPRMPAPRQVLOKFMKVPKSMKKEWIKYNLELLT 182  
 DB 162 IAQIVAAIRPKYKIKNUVLFFSTNSNRAFLRPPNPQPMTFVRRPESHSERDMVRHSVWMT 221  
 QY 183 TIGSPG-LDREKALDVRKSIERCLCEGTQORQALAAILOSQSRVGLLRRIAVPTLVISGA 241  
 DB 222 AVSGPGHLDIKGTTRATAEKYQORNFHPLXVSQQLTALASRSILRFTKQIRANTLVHGN 281  
 QY 242 EDPLLPYQCGRDIAHPIPGARFELIEGMGHDIPERHPIRLIELIAGH 288  
 DB 282 KDGLVAPNHGKILAKVISNARFVLVDGMGHDLENYYYPYINGLISEH 328

RESULT 4  
 US-08-935-263-10  
 ; Sequence 10, Application US/08935263A  
 ; Patent No. 6117669  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furuichi, Yasuhiro  
 ; APPLICANT: Hoshino, Tatsuo  
 ; APPLICANT: Kimura, Hitoshi  
 ; APPLICANT: Kiyasu, Tatsuya  
 ; APPLICANT: Nagahashi, Yoshie  
 ; TITLE OF INVENTION: BIOTIN BIOSYNTHETIC GENES  
 ; FILE REFERENCE: Biotin Genes  
 ; CURRENT APPLICATION NUMBER: US/08/935,263A  
 ; CURRENT FILING DATE: 1997-09-22  
 ; EARLIER APPLICATION NUMBER: EP 96115540.5  
 ; EARLIER FILING DATE: 1996-09-27  
 ; NUMBER OF SEQ ID NOS: 23  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO 10  
 ; LENGTH: 267  
 ; TYPE: PRP  
 ; ORGANISM: Kurthia sp.  
 US-08-935-263-10

Query Match 20.8%; Score 314.5; DB 2; Length 267;  
 Best Local Similarity 33.1%; Pred. No. 1.3e-25;  
 Matches 92; Conservative 43; Mismatches 104; Indels 39; Gaps 9;  
 ;  
 QY 1 MAQVKANGITLVEEQQHRRHPSMLLMGLGGQIDWPPEFIRGLAERGRVFCFNRDA 60  
 DB 1 MPFVNHDSLEYEV--HCGGDPDLLLIMGLGYNLSM-HRTVPTLAKR-FKVIIVFNRGV 56  
 QY 61 GLSTKLEGVKKPNIRARVFLASMLKPRVPTLDDMALDTVGLMDALGISTHVVGVSMG 120  
 DB 57 GKSS-----KPEQFYSIEMMAEDARAVLDVAVSDSAHVYGISMG 95  
 QY 121 GMAQILGAKHGRVSKLTMTTSSGNPRMPAPRQV---LQKFMKVPKSMKKEW--- 173  
 DB 96 GMAQRLATYPERVRSVLGCTTAGTTHIQPSPEISTLMVSRASLTGSPRDNALAAP 155  
 QY 174 IKYNLELLTIGSPGLDREKALDVRKSIERCLCEGTQORQALAAILOSQSRVGLLRRIAV 233  
 DB 156 IVYSQAFIEK--HPELIQE---DIQKRIEITTPSAVLSQLOACLTHDTSNE-LDKINI 208  
 QY 234 PTLVTSAGBDDPLLPYQCGRDIAHPIPGARFELIEGMGH 271  
 DB 209 PTLIIHGADNLVPYENGKMLAERIQQSQFHTVSCAGH 246

RESULT 5  
 US-09-594-185-10  
 ; Sequence 10, Application US/09594185  
 ; Patent No. 6365388  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furuichi, Yasuhiro  
 ; APPLICANT: Hoshino, Tatsuo  
 ; APPLICANT: Kimura, Hitoshi  
 ; APPLICANT: Kiyasu, Tatsuya  
 ; APPLICANT: Nagahashi, Yoshie

QY 1 MAQVKANGITLIEYEEQGRHHPMSLLIMGLGGQLIDWPPEEFIRGLAERGFVFCFDRDA 60



GENERAL INFORMATION:  
 APPLICANT: Marc J. Rubenfield et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 107196.136  
 CURRENT APPLICATION NUMBER: US/09/252,991A  
 CURRENT FILING DATE: 1999-02-18  
 PRIOR APPLICATION NUMBER: US 60/074,788  
 PRIOR FILING DATE: 1998-02-18  
 PRIOR APPLICATION NUMBER: US 60/094,190  
 PRIOR FILING DATE: 1998-07-27  
 NUMBER OF SEQ ID NOS: 33142  
 SEQ ID NO 32242  
 LENGTH: 268  
 TYPE: PRT  
 ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-32242

Query Match 12.5%; Score 189; DB 2; Length 268;  
 Best Local Similarity 25.5%; Pred. No. 5.3e-12;  
 Matches 75; Conservative 38; Mismatches 115; Indels 66; Gaps 10;

Qy 11 LEVEEQHRRHPSMLLGLGQGLDWPPEFIRGLAERGFVPCFDRDAGLSTKLGK 70  
 Db 16 LNSLEGAGAPVLLSLGSLGDLGMDTQ-IPALTAH-FRVRVYTRKHGASLVTG-- 71

Qy 71 KPNTARVELLASMGLKPRVPYTLDDMALDTVGLMDALGIESTHVVGSMGMIAQILGAK 130  
 Db 72 -----PVAIGQAGADVLALLDALELPRVHFGLSGMGLGQGLGIGH 112

Qy 131 HGERVKSJTLMTSS-----GNPRMPAPRPQVL---QKFMVPPKSMKDEWIKYNLEL 180  
 Db 113 AGERLGRVLNCTAAKIASDEWNTRIDT---VLKGGEQAMRVLRDASVARWF----- 162

Qy 181 LTTIGSPGLDREKALDVKSIERCLCEGTQRLAAILQSSRVKLLRRAVPTLVISG 240  
 Db 163 --TAGF--AEREPAQVRIVAMLAATSPOGVAANCAVRDADFR-EQGLVQAPTLIVAG 217

Qy 241 AEDPLPYQCCGRDIADHIPGARFELIEGMDHIDIPERHPIRLIETIAGHAAAEA 294  
 Db 218 SHDATTVPDNRPFQARIADA-----QLVEFAAHLNSVEA 253

RESULT 12  
 US-09-902-540-11639  
 Sequence 11639, Application US/09902540  
 Patent No. 6833447  
 GENERAL INFORMATION:  
 APPLICANT: Goldman, Barry S.  
 APPLICANT: Hinkle, Gregory J.  
 APPLICANT: Slater, Steven C.  
 APPLICANT: Wiegand, Roger C.  
 TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
 FILE REFERENCE: 38-10(15849)B  
 CURRENT APPLICATION NUMBER: US/09/902,540  
 CURRENT FILING DATE: 2001-07-10  
 PRIOR APPLICATION NUMBER: 60/217,883  
 PRIOR FILING DATE: 2000-07-10  
 NUMBER OF SEQ ID NOS: 16825  
 SEQ ID NO 11639  
 LENGTH: 265  
 TYPE: PRT  
 ORGANISM: Myxococcus xanthus  
 US-09-902-540-11639

Query Match 12.3%; Score 185; DB 2; Length 265;  
 Best Local Similarity 26.7%; Pred. No. 1.4e-11;  
 Matches 77; Conservative 38; Mismatches 121; Indels 52; Gaps 9;

Qy 1 MAQVANGITILEYEQHRRHPSMLLI--MGLGQGLDWPPEFIRGLAERGFVPCFDR 58  
 Db 1 MLTIVTGVPLHYRDEG--KGPPVLLHAFPLNGSTFD---KQVKAUSGR-YRFIIPDIR 54

Qy 59 DAGLSTKLEGVKKPNIAVRFVLLASMGKLPKRPVYTLDDMALDTVGLMDALGIESTHVVGVS 118  
 Db 55 GFGEALGDG-----PTMFRIARDALSLDLDALNDTIVVVGVS 93

Qy 119 MGGMIAQILGAKHGERVKSJTLMT-----SSGNPRMPAPRPQVLQKFMVPPKSMKDEE 172  
 Db 94 MGYAANALLREDAGRVSLVMDTQATADDAECKARRETSIAQALE-----VGVEP 145

Qy 173 WIKYNLELTTIGSPGLDREKALDVKSIERCLCEGTQRLAAILQSSRVKLLRRIA 232  
 Db 146 IIQAMLPKMWAAAGPDSVPAQEV-----ALMRAASPASVAAALRGMLRPSKDMRLARYA 200

Qy 233 VPTLVISGAEDPLPYQCCGRDIADHIPGARFELIEGMDHIDIPERHPIRLIETIAGH 276  
 Db 201 GPALVIVGEHDALTTPPAKAKEIAGLISGAKLEVIPDAGHLANQEPOR 248

RESULT 13  
 US-09-469-211A-6  
 Sequence 6, Application US/09469211A  
 Patent No. 6660524  
 GENERAL INFORMATION:  
 APPLICANT: J. Archer  
 TITLE OF INVENTION: CONTROL OF GENE EXPRESSION IN EUKARYOTES  
 FILE REFERENCE: 9341-021  
 CURRENT APPLICATION NUMBER: US/09/469,211A  
 CURRENT FILING DATE: 1999-12-22  
 PRIOR APPLICATION NUMBER: UK 9828660.2  
 PRIOR FILING DATE: 1998-12-24  
 NUMBER OF SEQ ID NOS: 19  
 SOFTWARE: Patent in Ver. 2.1  
 SEQ ID NO 6  
 LENGTH: 289  
 TYPE: PRT  
 ORGANISM: Rhodococcus sp.  
 US-09-469-211A-6

Query Match 12.1%; Score 183; DB 2; Length 289;  
 Best Local Similarity 24.3%; Pred. No. 2.7e-11;  
 Matches 68; Conservative 43; Mismatches 123; Indels 46; Gaps 9;

Qy 4 VKANGITILEYEQHRRHPSMLLGLGQGLDWPPEFIRGLAERG--FRVICDNRDAG 61  
 Db 22 IQAGPYTRYLHAGDSSKPTLLILHGITH-----AEAYVRNLRSHEFNWVADFIHG 77

Qy 62 LSTKLEGVKKPNIAVRFVLLASMGKLPKRPVYTLDDMALDTVGLMDALGIESTHVVGSMGG 121  
 Db 78 YST-----KPDHPIEKHYIDHVLQDLDAIGVEKASFSGESLGG 116

Qy 122 MIAQILGAKHGERVKSJTLMTSSGNPRMPAPRPQVLQKFMVPPKSMKDE---EWIKYNL 178  
 Db 117 WVTAFQAHDPKPKVDRIVLN-TWGGT---MANPQVMERLYTSLSWEAAKQPSWERSVKARL 171

Qy 179 ELL---TTIGSPGLDREKAL-----DVRKSIERCLCEGTQRLAAILQSSRVKLLRRI 231  
 Db 172 EWLMDPTMTVTDLIRTRQAIQQPDWLKACEMMALQDLETRKNMITDAT----LNGI 227

Qy 232 AVPTLVISGAEDPLPYQCCGRDIADHIPGARFELIEGMDHIDIPERHPIRLIETIAGH 271  
 Db 228 TVPAMVLWTTKDFSGPVDKRIASHIPGAKLAIMENCGH 267

RESULT 14  
 US-09-446-681-6  
 Sequence 6, Application US/09446681  
 Patent No. 6849442  
 GENERAL INFORMATION:  
 APPLICANT: Archer, John AC  
 APPLICANT: Summers, David K  
 APPLICANT: Roland, Herve J  
 APPLICANT: Powell, Justin AC  
 TITLE OF INVENTION: Biosensor materials and methods

Query Match 12.3%; Score 185; DB 2; Length 265;  
 Best Local Similarity 26.7%; Pred. No. 1.4e-11;  
 Matches 77; Conservative 38; Mismatches 121; Indels 52; Gaps 9;

Qy 1 MAQVANGITILEYEQHRRHPSMLLI--MGLGQGLDWPPEFIRGLAERGFVPCFDR 58  
 Db 1 MLTIVTGVPLHYRDEG--KGPPVLLHAFPLNGSTFD---KQVKAUSGR-YRFIIPDIR 54

```

RESULT 15
US-09-902-540-12298
; Sequence 12298, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12298
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12298

Query Match      12.1%; Score 182.5; DB 2; Length 271;
Best Local Similarity 26.8%; Pred. No. 2.7e-11;
Matches      84; Conservative 45; Mismatches 118; Indels 67; Gaps 14;

Q/      1  MAQVKANGITLVEEEOGHRHPSMLLIMGLGQLIDWPEEFIRGLAERGERVICFDNRDA 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1  MPLLQLEELSLYFEESSG--EGTPVLFLHGLGSSGRDN--ESVAPRLTGR-HRVIVPDARGH 56
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Q/      61  GLSTKLEGVKKPNIARVFLASMGKLPVPYFLDDMALDVTGLMDALGISTHVVGVSMD 120
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      57  GRSGKEPPGA-----YGVPRFARDIAGLCDALGLTGTHVVGLSMG 95
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 13, 2006, 13:10:48 ; Search time 161 Seconds  
(without alignments)  
762.993 Million cell updates/sec

Title: US-10-686-490D-2

Perfect score: 1509

Sequence: 1 MAQVKANGITLEYEEOGHRH.....ERHPLIELIAGHAAAAA 294

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA Main:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*

2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*

3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*

4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*

5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*

6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1509	100.0	294	5	US-10-686-490D-2
2	314.5	20.8	267	4	US-10-033-078-10
3	314.5	20.8	267	4	US-10-763-933-10
4	268.5	17.8	273	5	US-10-211-028-165
5	227.5	15.1	271	4	US-10-329-079-51
6	220.5	14.5	276	4	US-10-282-122A-49054
7	216.5	14.3	264	4	US-10-329-079-18
8	204	13.5	271	3	US-09-976-059-10
9	204	13.5	303	4	US-10-156-761-8055
10	198	13.1	232	3	US-09-896-578-4
11	198	13.1	232	3	US-09-971-490-14
12	198	13.1	232	4	US-10-289-148-4
13	198	13.1	232	4	US-10-038-854-66
14	198	13.1	232	4	US-10-193-452-36
15	198	13.1	232	4	US-10-193-452-87
16	197	13.1	217	4	US-10-051-874-95
17	193	12.8	224	4	US-10-099-322-303
18	193	12.8	224	4	US-10-044-564-303
19	191.5	12.7	272	5	US-10-211-028-8
20	189.5	12.6	251	4	US-10-068-134-4
21	185.5	12.3	233	4	US-10-056-744B-4
22	183.5	12.2	251	3	US-09-738-626-6126
23	183.5	12.2	251	5	US-10-494-675-50
24	183	12.1	289	5	US-10-732-859-6
25	181.5	12.0	264	4	US-10-724-972A-7349
26	180.5	12.0	258	4	US-10-282-122A-70509
27	180.5	12.0	310	4	US-10-425-115-292028

Sequence 78, Appl  
Sequence 43203, A  
Sequence 9245, Ap  
Sequence 11944, A  
Sequence 9087, Ap  
Sequence 46, Appl  
Sequence 45, Appl  
Sequence 65472, A  
Sequence 67314, A  
Sequence 2, Appl  
Sequence 8025, Ap  
Sequence 70255, A  
Sequence 66385, A  
Sequence 181547, A  
Sequence 9285, Ap  
Sequence 266316, A  
Sequence 54034, A  
Sequence 23, Appl

#### ALIGNMENTS

##### RESULT 1

US-10-686-490D-2

; Sequence 2, Application US/10686490D

; Publication No. US20050153404A1

; GENERAL INFORMATION:

; APPLICANT: Bayer Aktiengesellschaft

; TITLE OF INVENTION: Anti-Kazlauskas-Lipases

; FILE REFERENCE: Lea 35 991

; CURRENT APPLICATION NUMBER: US/10/686,490D

; CURRENT FILING DATE: 2003-10-15

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO: 2

; LENGTH: 294

; TYPE: PRT

; ORGANISM: unknown

; FEATURE:

; OTHER INFORMATION: bacterial

US-10-686-490D-2

Query Match 100.0%; Score 1509; DB 5; Length 294;

Best Local Similarity 100.0%; Pred. No. 3.6e-145;

Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAQVKANGITLEYEEOGHRHPSMLLMGGOLIDWPEEFIRGLAERGFVCFDNRDA	60
DB	1	MAQVKANGITLEYEEOGHRHPSMLLMGGOLIDWPEEFIRGLAERGFVCFDNRDA	60
QY	61	GLSTKLEGVKKNIA RVFLASMGKPRVPYPTLDDMALDTVGLMDALGIESTHVGVSMG	120
DB	61	GLSTKLEGVKKNIA RVFLASMGKPRVPYPTLDDMALDTVGLMDALGIESTHVGVSMG	120
QY	121	GMAIQLGAKHGSRVKS LTIMTSSGNPRMPAPRPOVLQKFMVPKSMDEEIKYNLEL	180
DB	121	GMAIQLGAKHGSRVKS LTIMTSSGNPRMPAPRPOVLQKFMVPKSMDEEIKYNLEL	180
QY	181	LTTIGSPGLDREKALDVKSTIERCLCPGTQQLAAILOSGRVKLLRRIAVPTLVISG	240
DB	181	LTTIGSPGLDREKALDVKSTIERCLCPGTQQLAAILOSGRVKLLRRIAVPTLVISG	240
QY	241	AEDPLLPYOCGRDIADHIFGARFELIEGMDHIDPERHIFELIAGHAAAAA	294
DB	241	AEDPLLPYOCGRDIADHIFGARFELIEGMDHIDPERHIFELIAGHAAAAA	294

##### RESULT 2

US-10-033-078-10

; Sequence 10, Application US/10033078

; Publication No. US20020123109A1

; GENERAL INFORMATION:



; APPLICANT: Furuichi, Yasuhiro  
 ; APPLICANT: Hoshino, Tatsuo  
 ; APPLICANT: Kimura, Hitoshi  
 ; APPLICANT: Kiyasu, Tatsuya  
 ; APPLICANT: Nagahashi, Yoshie  
 ; TITLE OF INVENTION: BIOTIN BIOSYNTHETIC GENES  
 ; FILE REFERENCE: Biotin Genes  
 ; CURRENT APPLICATION NUMBER: US/10/033,078  
 ; CURRENT FILING DATE: 2001-12-27  
 ; PRIOR APPLICATION NUMBER: 09/594,185  
 ; PRIOR FILING DATE: 2000-06-14  
 ; PRIOR APPLICATION NUMBER: EP 96115540.5  
 ; PRIOR FILING DATE: 1996-09-27  
 ; NUMBER OF SEQ ID NOS: 23  
 ; SOFTWARE: Patent in Ver. 2.1  
 ; SEQ ID NO 10  
 ; LENGTH: 267  
 ; TYPE: PRT  
 ; ORGANISM: Kurthia sp.  
 ; US-10-033-078-10

Query Match	20.8%	Score	314.5;	DB	4;	Length	267;		
Best Local Similarity	33.1%	Pred.	No. 2.9e-23;						
Matches	92;	Conservative	43;	Mismatches	104;	Indels	39;	Gaps	9;
Qy	1	MAQVKANGITLEYEEQGHRRHPSMLLIMGLOGQLDWPDEEPIRGLAERGFVIFCDNRDA	60						
Dd	1	MPFVNHDNESLYYEV--HGQGDPELLINGLVNSLSW-HRTVPTLAKR-FKVIIVDRGV	56						
Qy	61	GLSTKLGVKKPNIARVFLASMGILKRPVPTTDDMALDTVLGMDALGIESTHVUGVSMG	120						
Dd	57	GKSS-----KPEQPYSIEKMAEDARAVLDVASVDSAHVYGISMG	95						
Qy	121	GMTAQILGAKHGVRVKSITLMITSGGNPRMPPAPQV---LQKFMRVPKSDKEEW----	173						
Dd	96	GMTAQRLAITYPVRVSLVLGCTTAGGTTHIQSPSEISTLMVSRASSLTGSPRDNAWLAA	155						
Qy	174	IKNLELLTTIGSGLDREKLADVRKSIERCLOPEGTOROLAAIIQSGSRVKLLIRIAY	233						
Dd	156	IVTSQAPIEK--HPBLIQE-----DIQKRRIETTPSAVLSQLQACLTHDTSNE-LDKINI	208						
Qy	234	PTLVISAEDPPLLPGQCGRDIADHIPGARFELIEGMGH	271						
Dd	209	PTLIIGHGDAMLVYPYENGKMLAEIRIOGSOFHTVCAGH	246						

	Query Match	20.8%	Score 314.5;	DB 4;	Length 267;
	Best Local Similarity	33.1%;	Pred.No.2.9e-23;		
	Matches 92; Conservative	43;	Mismatches 104;	Indels 39;	Gaps 9;
Qy	1 MAQVKANGITLVEEEOGHRHPHSMLLIMGLGGOLLDPPEEFIRGLAERGRFVCFNRDA	60			
Dd	1 MPFNVDNESLYLEV--HGQGDPLLLMGLGYNSLSW-HRTVPTLAKR-FKVIYFDNRGV	56			
Qy	61 GLSTKLEGVKKPNIARVFILLASNGLPRVPDYTLDDNALDVTGLMDALGIESTHVGVSG	120			
Dd	57 GKSS-----RPEQYSIEMMAEDARVLDAVSVDASHVYGISMG	95			
Qy	121 GMTAIQILGAKHGSRVSKLTLMITS SGNPRMPAPRPQV---LQKFMRVPKSDKEEW---	173			
Dd	96 GMTIAQLAITYPERVRSVLGCTTAGTTHIQPSPEISTLMWSRASLTGSPRONAWLAAP	155			
Qy	174 IKYNLELLTTIGSPDLREKLADVRKSIEBCLCEPCTQPOLAAIILOGSRRVKLLRIAV	233			
Dd	156 IVTSQAFLIEK--HPELIQE----DIQRIRIITPPSAIYLSOLOACLTHDTISNE-LDKINI	208			
Qy	234 PTLVISGAEDPLLPYOCGRDIADIHIPGARFELIEGMGH	271			
Dd	209 PTLLIHGDADNLVPYENGKMLASRIQGSOFTHTVSCAGH	246			

Query Match	17.8%;	Score 268.5;	DB 5;	Length 273;
Best Local Similarity	26.0%;	Pred. No. 1.5e-18;		
Matches	85;	Conservative 44;	Mismatches 109;	Indels 79; Gaps 10;
Qy	1	MAQVKANGTLEVEEQ----	GHRHHPSMLLIMGLGGOLIDWPEEFIRGLAERFVRCIFD	56
Dd	1	MPVLTVNGIRINYYDDAPGAGNAQPAFLVMVGSGSGRAWHLHQVPALVAAGFRVISFD	60	
Qy	57	NRDAGLSTKLEGYKKENIARVFLLASMGKLPRVPYTTLDDMALDTGVGLMDALGSTHVVG	116	
Dd	61	NRCIAPSSECPG-----	GFGDIDLVDATAALVELRLGPCRVAG	99
Qy	117	VMSGMIQAQLNKAHGERSVKSLTMLITSSGNP-----	RNPAPRPQV	157
Dd	100	ISGAHIAOELALSRLPDVRLVLMATRA-RPDALREALCRAEMELYDQIRLPAAAYEAV	158	
Qy	158	LQKFMRV-PKSMDKE----EWIKYNLELTITGS	PGLDREKLADYRKSIERCLCPEGTQ	212
Dd	159	VQAMQLSPRTLNDVQARDWLDV-JELTRRSGA-----	GYR	194

QY 213 ROLAAILOSGRVKKLRRIAVPTLVISGAEDPLLPYQCGRDIAHPIGAFELIEGMGH- 271  
 Dp 195 AOL-GVRVDGRREAYGIRATRVAFQDDLIAPPHLGRVADAPGAEYELVPCDGHY 253  
 QY 272 ---DIPERHPIRLIELI 285  
 Db 254 GYLESFVAVKSLVEFL 270

RESULT 5  
 US-10-329-079-51  
 ; Sequence 51, Application US/10329079  
 ; Publication No. US20030198981A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FARNET, Chris  
 ; APPLICANT: ZAZOPOULOS, Emmanuel  
 ; APPLICANT: STAFFA, Alfredo  
 ; TITLE OF INVENTION: GENES AND PROTEINS INVOLVED IN THE BIOSYNTHESIS OF LIPOPEPTIDES  
 ; FILE REFERENCE: 3002-11US  
 ; CURRENT APPLICATION NUMBER: US/10/329,079  
 ; CURRENT FILING DATE: 2002-12-24  
 ; NUMBER OF SEQ ID NOS: 66  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 51  
 ; LENGTH: 271  
 ; TYPE: PRT  
 ; ORGANISM: Streptomyces refuineus  
 US-10-329-079-51

Query Match 15.1%; Score 227.5; DB 4; Length 271;  
 Best Local Similarity 26.9%; Pred. No. 2.2e-14;  
 Matches 84; Conservative 47; Mismatches 98; Indels 83; Gaps 13;

QY 1 MAQVKANGITLEYEEQGRHHPMSMLLMGLGQLIDWPPEEFIRGLAERGRVFCFNRDA 60  
 Db 1 MPTRINGIALDHRTG--SGPPVLLINGSAAKSAHMLHQVPALVAEGFEAVTFNR-- 56  
 QY 61 GLSTKLGKVPKPNIAVFLVLLASMGKPR---VPVTLDDMALDTVGLMDALGISTHVVGV 117  
 Db 57 -----GVPPSGGGGPGFTGLGMAADTVGLIHLGIGCAVVG 93  
 QY 118 SMGGMIAQILCAKGERVKSLTMTSSGNPRMA-----PRPVQLQK 160  
 Db 94 SLGARVAREARTPDLVRSCLVAPRARSRMACTAAEIALADSGVTLPPRYRAVR 153  
 QY 161 FMR--VPKSM--DKE--EWIKYNLELLTTIGSPGLDREKLALDVRSKIERCLCEGTORQ 214  
 Db 154 AMQNLSPRTLADRDQIADLDV-LELAADG-PGL-----RTQLE 191  
 QY 215 LAALQSGSRVKLRRIAVPTLVISGAEDPLLPYQCGRDIAHPIGAFELIEGMGH--- 271  
 Db 192 LSA---ADDRGEDLAGITAPCRVIAFADDIVAPPHLAKEIADALPEADYHVVPDCGHY 248  
 QY 272 -DIPERHPIRLI 282  
 Db 249 LERPDR-VNRLI 259

RESULT 6  
 US-10-282-122A-49054  
 ; Sequence 49054, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Malone, Cheryl  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari  
 ; APPLICANT: Zvekand, Judith  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John  
 ; APPLICANT: Carr, Grant  
 ; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.  
 ; APPLICANT: Xu, H.  
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 ; FILE REFERENCE: ELITRA.034A  
 ; CURRENT APPLICATION NUMBER: US/10/282.122A  
 ; CURRENT FILING DATE: 2003-02-20  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR FILING DATE: 2000-09-06  
 ; PRIOR FILING DATE: 2000-09-06  
 ; PRIOR FILING DATE: 2000-09-09  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR FILING DATE: 2001-02-16  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 78614  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 49054  
 ; LENGTH: 276  
 ; TYPE: PRT  
 ; ORGANISM: Burkholderia fungorum  
 US-10-282-122A-49054

Query Match 14.6%; Score 220.5; DB 4; Length 276;  
 Best Local Similarity 28.8%; Pred. No. 1.2e-13;  
 Matches 84; Conservative 43; Mismatches 104; Indels 63; Gaps 12;

QY 3 QVKANGITLEYEEQGRHHPMSMLLMGLGQLIDWPPEEFIRGLAERGRVFCFNRDAGL 62  
 Db 8 ETTVRCOTISYLELGDGPR-TLLLLHGITSASNWLST-MPALAQRGRVVIAPDLGFGQ 65  
 QY 63 STKLGKVPKPNIAVFLVLLASMGKPRVPVTLDDMALDTVGLMDALGISTHVVGVSGMG 122  
 Db 66 SS-----KPSV-----PVRPRTLSDM---VAPLLDALGLEMVSIVGQSMGGH 104  
 QY 123 IAQILCAKGERVKSLTMTTSSG--NPRMAPR-----PQVLOKEMRVKSMKE-- 171  
 Db 105 VAGLFAAQYPERVEALVNLVAGYGLALPEVKDPRDLGHAVTPGGL--WALNPATRDDSR 162  
 QY 172 --EWIKYNLELLTTIGSPGLDREKLALDVRSKIERCLCEGTORQLAAILQSGSR----- 224  
 Db 163 LLEMVFHDQGLVTEELIDGFYADRLG-----KGDGAVIRSISESWARREDTL 209  
 QY 225 ----VKLRRIAVPTLVISGAEDPLLPYQCGRDIAHPIGAFELIEGMGHDP 274  
 Db 210 ESAFTGLERR---PVLVIQARQDKVAPYHLGRAIHEGIAGSRFVLEDCGHAPP 260

RESULT 7  
 US-10-329-079-18  
 ; Sequence 18, Application US/10329079  
 ; Publication No. US20030198981A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FARNET, Chris  
 ; APPLICANT: ZAZOPOULOS, Emmanuel  
 ; APPLICANT: STAFFA, Alfredo  
 ; TITLE OF INVENTION: GENES AND PROTEINS INVOLVED IN THE BIOSYNTHESIS OF LIPOPEPTIDES  
 ; FILE REFERENCE: 3002-11US  
 ; CURRENT APPLICATION NUMBER: US/10/329,079  
 ; CURRENT FILING DATE: 2002-12-24  
 ; NUMBER OF SEQ ID NOS: 66



```
; Patent No. US20020077310A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 32225, A NOVEL HUMAN ALPHA BETA HYDROLASE
; TITLE OF INVENTION: FAMILY MEMBER
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 10448-073001
; CURRENT APPLICATION NUMBER: US/09/896,578
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/214,948
; PRIOR FILING DATE: 2000-06-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: consensus sequence
US-09-896-578-4

Query Match      13.1%; Score 198; DB 3; Length 232;
Best Local Similarity 31.5%; Pred. No. 1.9e-11;
Matches 82; Conservative 36; Mismatches 90; Indels 52; Gaps 12;

Qy 50 FRVICFNRDAGLSTKLEGVKKNIAKVFLVFLASGLKPRVPYTLDDMALDTVGLMDALGI 109
Db 1 FRVILLDLRGFGESSPSD-----LAE-----YRFDDLAEDLEALLDALGL 40
Qy 110 ESTHV-VGVSMGGMIAQILGAKHGE-RVKSLLTMITSSGNPRMPA-----PRPQVLQK 160
Db 41 EKPVLVGHSMGGAIALAYAAKYPELRVKALVLV-----SPPLPAGLSSDLFPRQGNLEG 95
Qy 161 FM-----RVPKSMDEXEIKYNLELLTTIGSPGL-DREKALDVRKSIERCLCPECTOR 213
Db 96 LLANFNRLRSV--EALLGRALKQFFLLGRPLVSDFLKQAEWLSLIRQGEDDGDG 153
Qy 214 QLAAILQSGSRVK-----LLRRIAVPTLVISGAEDPLLPYQCGRDIADHIFGARFELTEGM 269
Db 154 LLGNAVALGKLQWDLKXDKIKVPTLVITWGTDDPLVPLDASEKLSALIPNAEVVVIDDA 213
Qy 270 GH-----DIPERHIPRLIELI 285
Db 214 GHLALLEKPE-EVAELIKFL 232

RESULT 12
US-10-289-148-4
; Sequence 4, Application US/10289148
; Publication No. US2003010001A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Spurling, Heidi L.
; TITLE OF INVENTION: 46694, A HUMAN ALPHA/BETA HYDROLASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREFOR
; FILE REFERENCE: MP101-295PIRNM
; CURRENT APPLICATION NUMBER: US/10/289,148
; CURRENT FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: US 60/334,225
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Alpha/Beta Hydrolase Fold Consensus Domain
US-10-289-148-4

Query Match      13.1%; Score 198; DB 4; Length 232;
Best Local Similarity 31.5%; Pred. No. 1.9e-11;
Matches 82; Conservative 36; Mismatches 90; Indels 52; Gaps 12;

Qy 50 FRVICFNRDAGLSTKLEGVKKNIAKVFLVFLASGLKPRVPYTLDDMALDTVGLMDALGI 109
Db 1 FRVILLDLRGFGESSPSD-----LAE-----YRFDDLAEDLEALLDALGL 40
Qy 110 ESTHV-VGVSMGGMIAQILGAKHGE-RVKSLLTMITSSGNPRMPA-----PRPQVLQK 160
Db 41 EKPVLVGHSMGGAIALAYAAKYPELRVKALVLV-----SPPLPAGLSSDLFPRQGNLEG 95
Qy 161 FM-----RVPKSMDEXEIKYNLELLTTIGSPGL-DREKALDVRKSIERCLCPECTOR 213
Db 96 LLANFNRLRSV--EALLGRALKQFFLLGRPLVSDFLKQAEWLSLIRQGEDDGDG 153
Qy 214 QLAAILQSGSRVK-----LLRRIAVPTLVISGAEDPLLPYQCGRDIADHIFGARFELTEGM 269
Db 154 LLGNAVALGKLQWDLKXDKIKVPTLVITWGTDDPLVPLDASEKLSALIPNAEVVVIDDA 213
Qy 270 GH-----DIPERHIPRLIELI 285
Db 214 GHLALLEKPE-EVAELIKFL 232

; Patent No. US20020077310A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 32225, A NOVEL HUMAN ALPHA BETA HYDROLASE
; TITLE OF INVENTION: FAMILY MEMBER
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 10448-073001
; CURRENT APPLICATION NUMBER: US/09/896,578
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/214,948
; PRIOR FILING DATE: 2000-06-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: consensus sequence
US-09-896-578-4

Query Match      13.1%; Score 198; DB 3; Length 232;
Best Local Similarity 31.5%; Pred. No. 1.9e-11;
Matches 82; Conservative 36; Mismatches 90; Indels 52; Gaps 12;

Qy 50 FRVICFNRDAGLSTKLEGVKKNIAKVFLVFLASGLKPRVPYTLDDMALDTVGLMDALGI 109
Db 1 FRVILLDLRGFGESSPSD-----LAE-----YRFDDLAEDLEALLDALGL 40
Qy 110 ESTHV-VGVSMGGMIAQILGAKHGE-RVKSLLTMITSSGNPRMPA-----PRPQVLQK 160
Db 41 EKPVLVGHSMGGAIALAYAAKYPELRVKALVLV-----SPPLPAGLSSDLFPRQGNLEG 95
Qy 161 FM-----RVPKSMDEXEIKYNLELLTTIGSPGL-DREKALDVRKSIERCLCPECTOR 213
Db 96 LLANFNRLRSV--EALLGRALKQFFLLGRPLVSDFLKQAEWLSLIRQGEDDGDG 153
Qy 214 QLAAILQSGSRVK-----LLRRIAVPTLVISGAEDPLLPYQCGRDIADHIFGARFELTEGM 269
Db 154 LLGNAVALGKLQWDLKXDKIKVPTLVITWGTDDPLVPLDASEKLSALIPNAEVVVIDDA 213
Qy 270 GH-----DIPERHIPRLIELI 285
Db 214 GHLALLEKPE-EVAELIKFL 232

RESULT 11
US-09-971-490-14
; Sequence 14, Application US/09971490
; Publication No. US20030032091A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: 48120, 23479 AND 46689, NOVEL HUMAN HYDROLASES
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 10448-100001
; CURRENT APPLICATION NUMBER: US/09/971,490
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/238,170
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/237,991
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Consensus sequence
US-09-971-490-14
```



APPLICANT: Meyers, Rachel E.  
 APPLICANT: Glucksmann, Maria Alexandra  
 APPLICANT: Curtis, Rory A. J.  
 APPLICANT: Rudolph-Owen, Laura A.  
 TITLE OF INVENTION: NOVEL HUMAN HYDROLASE FAMILY MEMBERS AND USES THEREOF  
 FILE REFERENCE: 10448-194001  
 CURRENT APPLICATION NUMBER: US/10/193,452  
 PRIOR FILING DATE: 2002-07-11  
 PRIOR APPLICATION NUMBER: US 09/816,664  
 PRIOR FILING DATE: 2001-03-23  
 PRIOR APPLICATION NUMBER: US 60/191,973  
 PRIOR FILING DATE: 2000-03-24  
 PRIOR APPLICATION NUMBER: US 09/841,880  
 PRIOR FILING DATE: 2001-04-24  
 PRIOR APPLICATION NUMBER: US 60/199,559  
 PRIOR FILING DATE: 2000-04-25  
 PRIOR APPLICATION NUMBER: US 09/862,556  
 PRIOR FILING DATE: 2001-05-22  
 PRIOR APPLICATION NUMBER: PCT/US01/16424  
 PRIOR FILING DATE: 2001-05-22  
 PRIOR APPLICATION NUMBER: US 60/206,036  
 PRIOR FILING DATE: 2000-05-22  
 PRIOR APPLICATION NUMBER: US 09/861,165  
 PRIOR FILING DATE: 2001-05-18  
 PRIOR APPLICATION NUMBER: PCT/US01/16014  
 PRIOR FILING DATE: 2001-05-18  
 PRIOR APPLICATION NUMBER: US 60/205,442  
 PRIOR FILING DATE: 2000-05-19  
 PRIOR APPLICATION NUMBER: US 09/875,353  
 PRIOR FILING DATE: 2001-06-06  
 PRIOR APPLICATION NUMBER: PCT/US01/18335  
 PRIOR FILING DATE: 2001-06-06  
 PRIOR APPLICATION NUMBER: US 60/209,949  
 PRIOR FILING DATE: 2000-06-06  
 PRIOR APPLICATION NUMBER: US 09/896,578  
 PRIOR FILING DATE: 2001-06-29  
 PRIOR APPLICATION NUMBER: PCT/US01/20880  
 PRIOR FILING DATE: 2001-06-29  
 PRIOR APPLICATION NUMBER: US 60/214,948  
 PRIOR FILING DATE: 2000-06-29  
 PRIOR APPLICATION NUMBER: US 09/911,150  
 PRIOR FILING DATE: 2001-07-23  
 PRIOR APPLICATION NUMBER: PCT/US01/23153  
 PRIOR FILING DATE: 2001-07-23  
 PRIOR APPLICATION NUMBER: US 60/220,008  
 PRIOR FILING DATE: 2000-07-21  
 PRIOR APPLICATION NUMBER: US 09/911,317  
 PRIOR FILING DATE: 2001-07-23  
 PRIOR APPLICATION NUMBER: PCT/US01/23160  
 PRIOR FILING DATE: 2001-07-23  
 PRIOR APPLICATION NUMBER: US 60/220,040  
 PRIOR FILING DATE: 2000-07-21  
 PRIOR APPLICATION NUMBER: US 09/934,323  
 PRIOR FILING DATE: 2001-08-21  
 PRIOR APPLICATION NUMBER: PCT/US01/26091  
 PRIOR FILING DATE: 2001-08-21  
 PRIOR APPLICATION NUMBER: US 60/226,774  
 PRIOR FILING DATE: 2000-08-21  
 PRIOR APPLICATION NUMBER: US 09/963,959  
 PRIOR FILING DATE: 2001-09-25  
 PRIOR APPLICATION NUMBER: PCT/US01/29962  
 PRIOR FILING DATE: 2001-09-25  
 PRIOR APPLICATION NUMBER: US 60/235,033  
 PRIOR FILING DATE: 2000-09-25  
 PRIOR APPLICATION NUMBER: US 09/971,490  
 PRIOR FILING DATE: 2001-10-05  
 PRIOR APPLICATION NUMBER: PCT/US01/31674  
 PRIOR FILING DATE: 2001-10-05  
 PRIOR APPLICATION NUMBER: US 60/238,170  
 PRIOR FILING DATE: 2000-10-05  
 PRIOR APPLICATION NUMBER: US 10/071,275  
 PRIOR FILING DATE: 2002-02-07  
 PRIOR APPLICATION NUMBER: PCT/US02/03793

---

PRIOR FILING DATE: 2001-07-23  
 PRIOR APPLICATION NUMBER: US 60/220,040  
 PRIOR FILING DATE: 2000-07-21  
 PRIOR APPLICATION NUMBER: US 09/934,323  
 PRIOR FILING DATE: 2001-08-21  
 PRIOR APPLICATION NUMBER: PCT/US01/26091  
 PRIOR FILING DATE: 2001-08-21  
 PRIOR APPLICATION NUMBER: US 60/226,774  
 PRIOR FILING DATE: 2000-08-21  
 PRIOR APPLICATION NUMBER: US 09/963,959  
 PRIOR FILING DATE: 2001-09-25  
 PRIOR APPLICATION NUMBER: PCT/US01/29962  
 PRIOR FILING DATE: 2001-09-25  
 PRIOR APPLICATION NUMBER: US 60/235,033  
 PRIOR FILING DATE: 2000-09-25  
 PRIOR APPLICATION NUMBER: US 09/971,490  
 PRIOR FILING DATE: 2001-10-05  
 PRIOR APPLICATION NUMBER: PCT/US01/31674  
 PRIOR FILING DATE: 2001-10-05  
 PRIOR APPLICATION NUMBER: US 60/238,170  
 PRIOR FILING DATE: 2000-10-05  
 PRIOR APPLICATION NUMBER: US 10/071,275  
 PRIOR FILING DATE: 2002-02-07  
 PRIOR APPLICATION NUMBER: PCT/US02/03793  
 PRIOR FILING DATE: 2002-02-07  
 PRIOR APPLICATION NUMBER: US 60/267,054  
 PRIOR FILING DATE: 2001-02-07  
 PRIOR APPLICATION NUMBER: US 09/888,911  
 PRIOR FILING DATE: 2001-06-25  
 PRIOR APPLICATION NUMBER: PCT/US01/19967  
 PRIOR FILING DATE: 2001-06-25  
 PRIOR APPLICATION NUMBER: US 60/213,688  
 PRIOR FILING DATE: 2000-06-23  
 NUMBER OF SEQ ID NOS: 104  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 36  
 LENGTH: 232  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: consensus sequence  
 US-10-193-452-36

Query Match 13.1%; Score 198; DB 4; Length 232;  
 Best Local Similarity 31.5%; Pred. No. 1.9e-11;  
 Matches 82; Conservative 36; Mismatches 90; Indels 52; Gaps 12;

QY	50	FRVICFNDKAGTSLKLGKKENIARVFLLASMKGLKPRVPTLDDMALDTVLMDALGI	109
DB	1	FRVILLDLRGFGSSPSD-----LAE-----YRFDLAEDLEALLDALGL	40
QY	110	ESTHV-VGVSMGMAIQILGAHGE-RVKSLTLMTITSGNPRMPA-----PRPQVLOK	160
DB	41	EKPVIIVGHSMGMAIALVAAYKAPELRVALVV-----SPPLPAGLSDDLPPRQNLGG	95
QY	161	FM-----RVPKSMDEMIKYNNLELTITIGSEGL-DREKALDVRSKIERCICPEGTOR	213
DB	96	LLANFNRLSRV--EALLGRALKQFFLLGRPLVSDFLKQAEDWLSLIHQEGDDGGDG	153
QY	214	QLAAIIOSGRVK----LIIRIAVPTLVISGABDPLLPOCGRDIAHDHPGARFELIEGM	269
DB	154	LIGAVALGKLQWLDSALKDKIKVPTLVINGTDDPLVPLDASEKLSALIPNAEVVVVIDA	213
QY	270	GH-----DIPERHIPRIELI	285
DB	214	GHLALLEKEPE-EVAELIKFL	232

RESULT 15  
 US-10-193-452-87  
 ; Sequence 87, Application US/10193452  
 ; Publication No. US20040214758A1  
 ; GENERAL INFORMATION:

;; PRIOR FILING DATE: 2002-02-07  
;; PRIOR APPLICATION NUMBER: US 60/267,054  
;; PRIOR FILING DATE: 2001-02-07  
;; PRIOR APPLICATION NUMBER: US 09/888,911  
;; PRIOR FILING DATE: 2001-06-25  
;; PRIOR APPLICATION NUMBER: PCT/US01/19967  
;; PRIOR FILING DATE: 2001-06-25  
;; PRIOR APPLICATION NUMBER: US 60/213,688  
;; PRIOR FILING DATE: 2000-06-23  
;; NUMBER OF SEQ ID NOS: 104  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 87  
;; LENGTH: 232  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: consensus sequence  
US-10-193-452-87

Query Match 13.1%; Score 198; DB 4; Length 232;  
Best Local Similarity 31.5%; Pred. No. 1.9e-11;  
Matches 82; Conservative 36; Mismatches 90; Indels 52; Gaps 12;  
Qy 50 FRVICFNRDAGLSTKLEGVKKPNIRVFLASMGKLPKRVPTLDDMALDTVGLMDALGI 109  
Db 1 FRVILLDLRGFGESSPSD-----LAE-----YRFDLAEDELLDALGL 40  
Qy 110 ESTHV-VGVSMGGMIAQILGAKHGE-RVKSLLTMITSSGNRMPA-----PRQVLOK 160  
Db 41 EKPVLVGHSMGGAIALAYAAKYPRLRYKALV-----SPPLPAGLSSDLFPRQGNLEG 95  
Qy 161 FM-----RVPKSMCKEWMKYNLELLTTIGSPGL-DREKALDVRKSIERCLCEGTQR 213  
Db 96 LILANPRNLSRSV--EALLGRALKQFLLGRPLVSDFLKQAEWLSLIRQGEDDGGD 153  
Qy 214 QLAAILQSGSRVK----LLRRIAVPTLVISGAEDPLLPYQCGRDIADHIPGARFELIEGM 269  
Db 154 LLGAVALGKLLQWDLGALKDKVPTLVINGTDDPLVPLDASEKLSALIPNAEVVVWIDDA 213  
Qy 270 GH-----DIPERHPIRLIELI 285  
Db 214 GHLELEKPE-EVAELIKFL 232

Search completed: April 13, 2006, 13:14:23  
Job time : 162 secs



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Ran on: April 13, 2006, 13:11:49 ; Search time 27 Seconds  
(without alignments)  
462.661 Million cell updates/sec

Title: US-10-686-490d-2  
Perfect score: 1509  
Sequence: 1 MAQVANGITLVEEEOGHRH.....ERHPRILIELIAGHAAAEA 294

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 217505 seqs, 42489236 residues

Total number of hits satisfying chosen parameters: 217505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:

1: /SID55/ptodata/1/pubpaa/US08\_NEW\_PUB pep.\*  
2: /SID55/ptodata/1/pubpaa/US06\_NEW\_PUB pep.\*  
3: /SID55/ptodata/1/pubpaa/US07\_NEW\_PUB pep.\*  
4: /SID55/ptodata/1/pubpaa/PCT\_NEW\_PUB pep.\*  
5: /SID55/ptodata/1/pubpaa/US03\_NEW\_PUB pep.\*  
6: /SID55/ptodata/1/pubpaa/US10\_NEW\_PUB pep.\*  
7: /SID55/ptodata/1/pubpaa/US11\_NEW\_PUB pep.\*  
8: /SID55/ptodata/1/pubpaa/US60\_NEW\_PUB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	209.5	13.9	260	7	US-11-087-099-1714 Sequence 1714, Ap
2	204	13.5	271	7	US-11-205-109-10 Sequence 10, Appl
3	201.5	13.4	312	7	US-11-087-099-5676 Sequence 5676, Ap
4	193	12.8	224	7	US-11-054-281-303 Sequence 303, App
5	187	12.4	265	7	US-11-087-099-4437 Sequence 4437, Ap
6	183.5	12.2	251	7	US-11-055-822-470 Sequence 470, App
7	183	12.1	265	7	US-11-087-099-9057 Sequence 9057, Ap
8	176.5	11.7	266	7	US-11-087-099-967 Sequence 967, App
9	176.5	11.7	266	7	US-11-087-099-8501 Sequence 8501, Ap
10	168.5	11.2	293	6	US-10-467-657-6664 Sequence 6664, Ap
11	160	10.6	2197	7	US-11-075-185-8 Sequence 8, Appli
12	155.5	10.3	286	7	US-11-179-977-9 Sequence 9, Appli
13	147	9.7	308	7	US-11-096-568A-25601 Sequence 25601, A
14	147	9.7	330	7	US-11-096-568A-25600 Sequence 25600, A
15	145.5	9.6	264	7	US-11-096-568A-29097 Sequence 29097, A
16	145.5	9.6	307	7	US-11-096-568A-29096 Sequence 29096, A
17	145.5	9.6	311	7	US-11-096-568A-29095 Sequence 29095, A
18	141.5	9.4	256	7	US-11-096-568A-29086 Sequence 29086, A
19	141.5	9.4	267	7	US-11-087-099-11146 Sequence 11146, A
20	141.5	9.4	299	7	US-11-096-568A-29085 Sequence 29085, A
21	141.5	9.4	303	7	US-11-096-568A-29084 Sequence 29084, A
22	136.5	9.0	289	7	US-11-179-977-11 Sequence 11, Appl
23	136.5	9.0	289	7	US-11-087-099-8151 Sequence 8151, Ap
24	135	8.9	251	7	US-11-096-568A-4794 Sequence 4794, Ap
25	130.5	8.6	256	7	US-11-179-977-15 Sequence 15, Appl

26	128	8.5	250	7	US-11-096-568A-33303 Sequence 33303, A
27	128	8.5	273	7	US-11-096-568A-33302 Sequence 33302, A
28	128	8.5	282	7	US-11-096-568A-33301 Sequence 33301, A
29	124	8.2	267	7	US-11-096-568A-4795 Sequence 4795, Ap
30	122	8.1	331	6	US-10-454-437-226 Sequence 226, App
31	121.5	8.1	302	7	US-11-087-099-10293 Sequence 10293, A
32	120.5	8.0	311	6	US-10-467-657-5644 Sequence 5644, Ap
33	120.5	8.0	334	6	US-10-525-674-30 Sequence 30, Appl
34	119	7.9	293	7	US-11-096-568A-25602 Sequence 25602, A
35	117.5	7.8	487	7	US-11-096-568A-13018 Sequence 13018, A
36	117.5	7.8	546	7	US-11-096-568A-13017 Sequence 13017, A
37	117	7.8	284	7	US-11-096-568A-5989 Sequence 5989, Ap
38	117	7.8	401	7	US-11-096-568A-5988 Sequence 5988, Ap
39	117	7.8	416	7	US-11-096-568A-5987 Sequence 5987, Ap
40	114.5	7.6	311	7	US-11-096-568A-6768 Sequence 6768, Ap
41	113	7.5	324	7	US-11-096-568A-33701 Sequence 33701, A
42	113	7.5	380	6	US-10-525-674-28 Sequence 28, Appl
43	111.5	7.4	259	7	US-11-087-099-6077 Sequence 6077, Ap
44	110.5	7.3	269	7	US-11-087-099-9507 Sequence 9507, Ap
45	110.5	7.3	399	7	US-11-096-568A-21969 Sequence 21969, A

## ALIGNMENTS

## RESULT 1

US-11-087-099-1714  
; Sequence 1714, Application US/11087099  
; Publication No. US20060041961A1  
; GENERAL INFORMATION:  
; APPLICANT: Abad, Mark S. et al.  
; TITLE OF INVENTION: Genes and Uses for Plant Improvement  
; FILE REFERENCE: 38-21(53450)B EP  
; CURRENT APPLICATION NUMBER: US/11/087,099  
; CURRENT FILING DATE: 2005-03-22  
; NUMBER OF SEQ ID NOS: 12464  
; SEQ ID NO 1714  
; LENGTH: 260  
; TYPE: PRT  
; ORGANISM: Bradyrhizobium japonicum USDA 110  
US-11-087-099-1714

Query Match 13.9%; Score 209.5; DB 7; Length 260;

Best Local Similarity 26.1%; Pred. No. 4.1e-12;

Matches 74; Conservative 37; Mismatches 109; Indels 63; Gaps 8;

QY	1	MAQVANGITLVEEEOGHRHPSMLLMGGLIDMPREFIRGLAERGFVFCFNRDA	60
DB	1	MPMIDADGCLINVSVEGRDGGPTLMLNSLGCTLQMW-EPQMKALTO-VPRVIRYDRGH	58
QY	61	GLSTKLEGVKKNIAARVFLASMGKLPVPVYTLDDMALDTVGLMDALGIESTHVGVSMG	120
DB	59	G-----KSNV-----PPGPTWMEFRDVLAILNDLNIKVKHVGCLSMG	97
QY	121	GMAQILKAHGERVKSITLMTSSGNRMPAPRPOVLQKFMKVPKSMKKEIKYNLEL	180
DB	98	GMVGWGLGANAPERFGKLILANTSC-----YYAETPKV-----LER	133
QY	181	LTTIGSPGL-----DREKLADVRKSTERCICPEGTQRLAA---ILQSGSR	224
DB	134	IDAVKKGGAADVAVIAGWLQDFRFRPDITARMKAMLLASPVGYLACCALSTLDQ	193
QY	225	VKLLRRIAVPTLVISGAEDPLLPYQCGRDIADHIFGARGFELIE	267
DB	194	RELLPKIKSPTLVIAGRHDMATPISAGELIRSKIFGANNTIID	236

## RESULT 2

US-11-205-109-10  
; Sequence 10, Application US/11205109  
; Publication No. US20050287641A1  
; GENERAL INFORMATION:  
; APPLICANT: Parnet, Chris

APPLICANT: Zazopoulos, Emmanuel  
 APPLICANT: Staffa, Alfredo  
 TITLE OF INVENTION: GENE CLUSTER FOR RAMOPLANIN BIOSYNTHESIS

FILE REFERENCE: 3002-ZUS  
 CURRENT APPLICATION NUMBER: US/11/205,109

CURRENT FILING DATE: 2005-08-17

PRIOR APPLICATION NUMBER: US/09/976,059

PRIOR FILING DATE: 2001-10-15

PRIOR APPLICATION NUMBER: US 60/239,924

PRIOR FILING DATE: 2000-10-13

NUMBER OF SEQ ID NOS: 46

SOFTWARE: PatentIn version 3.0

SEQ ID NO 10

LENGTH: 271

TYPE: PRT

ORGANISM: Actinoplanes sp.

NAME/KEY: misc feature

LOCATION: (1)-(1)

OTHER INFORMATION: V represents a non-standard initiator codon. It is expected that the biosynthesized protein will have a formylmethionine residue at this position

US-11-205-109-10

Query Match 13.5%; Score 204; DB 7; Length 271;  
 Best Local Similarity 25.0%; Pred. No. 1.4e-11;  
 Matches 78; Conservative 38; Mismatches 120; Indels 76; Gaps 9;

Qy 4 VKANGITLEYEEOGHRHPSMLLIMGLGQGLIDWPPEFIRGLAERGRVFCFNRDAGLS 63

Db 9 VTTNGVRLAYRRSGAGE--PVLIMMGSSGAGQTVTQTPALHAGYSTVVFDSR----- 61

Qy 64 TKLEGVKKPNIRAVFLLASMLKPRVPYTLDDMALDTVGLMDALGIESTHVGVSMGMI 123

Db 62 ----GIPSDV-----PAGKYSLADMTADTGLIEALDLAPCRIVGTSLGAMI 105

Qy 124 AQILGAKHGERVKSLTMITSSGNPRMPAPRQ---VLQKFMVPKSMDK----- 170

Db 106 AQELAVDPHELVRCAVLIATLRPDARAQAQNDILLSGVTLPAAEYAAATAVFKMFS 165

Qy 171 -----EEWIKYNLELTITGSPGLDREKLDVRSIERCLCPEGTORQLAAILQ 220

Db 166 PATLNDVDVAVREWLDI-FELSGTGVSA-----GQAWAEL- 199

Qy 221 SGRVKLLRRIAVPTLVISGADPLLPYOCGRDIADHIPGARFELIEGMGH----- 276

Db 200 TGDRAALRSVTPACRVISFADLLITPPHAAEVAEAIPOCDLVEISRCGHGLYLERPDA 259

Qy 277 HIPRLIELIAGH 288

Db 260 VNAALFELDSH 271

RESULT 3

US-11-087-099-5676

Sequence 5676, Application US/11087099

Publication No. US20060041961A1

GENERAL INFORMATION:

APPLICANT: Abad, Mark S. et al.

TITLE OF INVENTION: Genes and Uses for Plant Improvement

FILE REFERENCE: 38-21(53450)B EP

CURRENT APPLICATION NUMBER: US/11/087,099

CURRENT FILING DATE: 2005-03-22

NUMBER OF SEQ ID NOS: 12464

SEQ ID NO 5676

LENGTH: 312

TYPE: PRT

ORGANISM: Rhodobacter sphaeroides

US-11-087-099-5676

Query Match 13.4%; Score 201.5; DB 7; Length 312;

Best Local Similarity 24.4%; Pred. No. 3e-11;

Matches 71; Conservative 42; Mismatches 105; Indels 73; Gaps 8;

Qy 1 MAQVKANGITLEYEEOGHRHPSMLLIMGLGQGLIDWPPEFIRGLAERGRVFCFNRDA 60  
 Db 51 MPHIEIGDURLNYLDEGRATGAPVVFHAHAGSDLRW--DNLLPLPDQLRLVRYDLRGH 108  
 Qy 61 GLSTKLEGVKKPNIRAVFLLASMLKPRVPYTLDDMALDTVGLMDALGIESTHVGVSMG 120  
 Db 109 GRST-----TPEPPYANGALIRDAERLMEALSIREAVFVGCSIG 147  
 Qy 121 GMTAQILGAKHGERVKSLTMITSSGNPRMPAPRQVQLQKFMVPKSMDKEEWIK-YNLE 179  
 Db 148 GMTAQGLAVKRLDLVRGLVLCDTAA-----KIGTPEIWQ-----DRIDQVRSYGLE 193  
 Qy 180 LTTIGSPGLDREKLDVRSIERCLCP---EGTQRL-----RAILQ 220  
 Db 194 SLAD-----PTMKRWAFAPFQGPGLWRERFIEGDDPGYAGGAAIA 237  
 Qy 221 SGRVKLLRRIAVPTLVISGADPLLPYOCGRDIADHIPGARFELIEGMGH 271  
 Db 238 GTDFYTTTARLTPTLALVSGEDGSTPPDLVRETASLIRGSRFEIVRGACH 288

RESULT 4

US-11-054-281-303

Sequence 303, Application US/11054281

Publication No. US20060013813A1

GENERAL INFORMATION:

APPLICANT: Mezes et al.

TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

FILE REFERENCE: 21402-240CIP

CURRENT APPLICATION NUMBER: US/11/054,281

CURRENT FILING DATE: 2005-02-08

PRIOR APPLICATION NUMBER: 60/261,014

PRIOR FILING DATE: 2001-01-11

PRIOR APPLICATION NUMBER: 60/261,018

PRIOR FILING DATE: 2001-01-11

PRIOR APPLICATION NUMBER: 60/318,410

PRIOR FILING DATE: 2001-09-10

PRIOR APPLICATION NUMBER: 60/261,013

PRIOR FILING DATE: 2001-01-11

PRIOR APPLICATION NUMBER: 60/261,026

PRIOR FILING DATE: 2001-01-11

PRIOR APPLICATION NUMBER: 60/261,029

PRIOR FILING DATE: 2001-01-11

PRIOR APPLICATION NUMBER: 60/313,170

PRIOR FILING DATE: 2001-08-17

PRIOR APPLICATION NUMBER: 10/044,564

PRIOR FILING DATE: 2002-01-11

NUMBER OF SEQ ID NOS: 324

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 303

LENGTH: 224

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence:consensus

OTHER INFORMATION: sequence

US-11-054-281-303

Query Match 12.8%; Score 193; DB 7; Length 224;

Best Local Similarity 28.1%; Pred. No. 1.2e-10;

Matches 71; Conservative 31; Mismatches 71; Indels 80; Gaps 9;

Qy 52 VICFDRDAGLSTKLEGVKKPNIRAVFLLASMLKPRVPYTLDDMALDTVGLMDALGIES 111  
 Db 1 VILFDLRGFGQSSPSD-----LAE-----YRPDDLAEADLEALLDGLDX 40

Qy 112 THVGVSMGGMIAQILGAKHGERVKSLTMITSSGNPRMPAPRQVQLQKFMVPKSM--- 168  
 Db 41 VILVGHSMGGAIAAAYAAKTPKRVKALV-----VSAPHPALLSSRL-FPRNFG 90

Qy 169 -----DKEEWIKYNLELTITGSPGLDREKLDVRSIERCLCPEGTORQLAA 217  
 Db 169 -----DKEEWIKYNLELTITGSPGLDREKLDVRSIERCLCPEGTORQLAA 217







